The British Journal for the
Philosophy of Science

VOLUME II MAY, 1951 No. 5

THE HYPOTHESIS OF CYBERNETICS

1 *Introduction.* 2 *Feed-Back.* 3 *The Basic Cybernetic Hypothesis.* 4 *Application and Confirmation.* 5 *Open Systems.* 6 *Electro-encephalography.* 7 *Evaluation.*

1 *Introduction. Automatic Calculating Machines*

DURING World War II some of the research in theoretical physics and engineering, which was important both for attack and defence, required vast calculations to be carried out. The delay involved in doing these with pencil and paper would have been a serious drawback; but the difficulty was overcome by means of electronic calculating machines, which could solve in a matter of hours mathematical equations that would otherwise have taken months to work out. Such machines had to be able (*a*) to be fed with information, (*b*) to store in some form the results of their preliminary calculations, (*c*) to make use of these results appropriately at the right time, i.e. to select from their ' files ', and (*d*) to report their final results. There is an obvious analogy between such powers and those of the intellect—working out, storing, recalling, selecting, and communicating. Indeed it is possible to design a machine to play chess without breaking the rules and even to make ' sound ' moves, once it has been provided with some criterion of evaluation of moves. If this were all, however, the analogy between the machine and the brain or mind (we need not here be precise which) would hardly be striking. What is important is that these machines embody a special type of mechanism.

2 *Feed-Back*

2.0 *Examples of Feed-Back.* These electronic machines are not merely 'automatic', in the sense in which a seesaw is automatic. They embody a mechanism known as "feed-back", which is of the highest importance; they thus differ from purely Newtonian machines (the nature of this difference requires elucidation and will be discussed in the final section).

A thermostat affords one very simple example of feed-back in everyday use. When the temperature rises too high, the extra heat operates in such a way as to cut off the source of heat, which reduces the temperature. Suppose an incubator, heated by gas, has a portion of gas pipe passing through it and that a tiny inflated balloon lies in the pipe. When the incubator gets hotter, the balloon expands and thus reduces the amount of gas passing through the pipe, perhaps even cutting off the supply. The burner of the incubator now receives less gas and cools. This makes the balloon contract, so that more gas can flow and once more warm up the incubator. It is to be observed that the temperature of the incubator is not absolutely constant; it moves very slightly above and below the mean temperature, making small errors which it then proceeds to correct.

Another example is to be found in a spring-driven gramophone. Governors, which are weights at the end of wires hinged to the spindle, will be found under the turn-table. When this rotates too fast, the weights swing away from the spindle, and this slows down the motion; the weights then fall in towards the spindle, which then rotates a little faster. Thus the record is nearly always a minute amount out of tune—sometimes very slightly sharp, sometimes very slightly flat, and just right in between. The literature of the subject generally cites the governors of Watt's steam-engine as providing one of the oldest mechanisms of this kind and credits Maxwell[1] in 1868 with giving the first theoretical account of them. But a very old example can be inspected, for instance, in the Science Museum in London, where one of the oldest mechanical clocks in the world, dating from the fourteenth century (unless it is a seventeenth-century model after the fourteenth-century manner), shows the mechanism in its ingenious escapement—in principle almost the same as the escapement in the modern watch. Rods forming the clock are rotated by means of a weight and a coil of rope. To prevent the weight crashing to earth and driving the clock at ever-increasing speed, a cog-wheel on one of the rods is caught by a lug attached to a spindle. The cog throws out the lug, which swings round its spindle and thereby swings another lug into another tooth of the cog-wheel.[1] The initial motion of the clock is thus checked. The process is then repeated indefinitely. Clearly the clock does not keep exact time. It is nearly always going slightly too fast or slightly too slow—only in between is it just right.

---

[1] J. C. Maxwell, "On Governors", *Proc. Roy. Soc. London*, 1868, **16**, 270

Riding a bicycle affords another example. When the rider falls slightly to the right, he turns his front wheel to the right, which stops his fall but leads to his being gently thrown over on the left. This he corrects by turning his wheel to his left.

2.1 *Feed-Back, Positive and Negative.* These are all examples of ' negative feed-back '. In all of them the work done by a source of power is controlled—the machine is prevented from running away—by a mechanism that checks it and keeps the rate at which energy is expended hovering round a certain quantity. This feed-back is ' negative ' because the work done by the feed-back mechanism opposes the working of the main driving force.

Positive feed-back is also possible. Suppose, for instance, that instead of having a balloon in the gas pipe of the incubator, we have one outside, fixed in such a way that, as it expands through being heated, it pushes a lever of a tap, thus partially opening the tap. This will increase the heat of the incubator, and warm the balloon still more, so that the tap will be opened further ; and so the process will continue.

Here, not only does the incubator become hotter simply through the consumption of a steady stream of gas, but a ' positive ' feed-back mechanism *increases the flow* of the gas. This may seem to be a futile type of mechanism—for why not remove the balloon, open the tap, and have the maximum flow of gas from the outset ? The reason is that there are circumstances when it is desirable to build up some required state at a slow pace. Suppose, for instance, we wish to run an electric train without a driver ; we must not turn on full power at once ; a positive feed-back must therefore be used to keep turning on more power gradually.

In all these mechanisms the energy of the feed-back comes from the given source of power driving the machine ; the negative feed-back mechanisms—governors, escapements, thermostats—are simply devices for re-directing some of the energy supplied by the source of power against the rest of that energy, and the positive feed-back mechanisms redirect some of the energy towards increasing the energy supplied by the source of power—towards removing a blockage in fact. Both negative and positive types here described may be referred to as " simple feed-back mechanisms ".

---

[1] What is known as the " angular momentum " of the escapement is adjusted very simply in this clock, but the details need not be considered here.

2.2 *Amplification Feed-Back.* In none of the foregoing examples does a local source of energy arising from the feed-back mechanism play a part. The energy of the feed-back is derivative and redirected. But an additional source of energy may be exploited and utilised in the same sort of way. For example, in the combustion of coal, the lighted paper ignites some coal which releases vastly greater energy which fires more coal and so on. Again, in the uranium bomb, a neutron splits an atomic nucleus which releases energy and more neutrons which split more atomic nuclei and release more energy. Any chain-reaction mechanism is an illustration. Another example is the inflationary spiral of increased wages leading to increased prices leading to increased wages. If we consider two human beings, the regard or animosity of one for the other may build up through positive feed-back—or it may become damped or allayed by negative feed-back. The additional source of energy may conveniently be spoken of as " local ", as contrasted with the main or central source.

Here again we have negative or positive feed-back according to whether the locally tapped energy is fed back to impede or facilitate the main motion of the machine. When the feed-back is due in any measure to the exploitation of local energy we may speak of " amplification ".

The energy of a feed-back may be solely derived from the main power or solely due to a local source, or to both, and it may be desirable to distinguish the regulating energy and mechanism as " simple feed-back ", " amplification feed-back ", and " mixed feed-back ". These are not well-distinguished in the literature, but in the field of cybernetics it is simple feed-back—and mainly negative at that—that writers usually have in mind. But this restriction is not required, and in what follows " feed-back " will be used indifferently for any form of the phenomenon ; where determinateness is needed the appropriate qualifying word will be added.

3 *The Basic Cybernetic Hypothesis*

We may describe machines that embody negative feed-back mechanisms as proceeding by " trial and error ", or as " error-compensating ", or—best—as " self-correcting " ; and we may define " a simple negative feed-back mechanism " as a mechanism by which part of the input-energy of a machine is utilised at intervals to impose a check on the output-energy.

The basic hypothesis of cybernetics is that the chief mechanism of the central nervous system is one of negative feed-back. The field of study is not, however, restricted to feed-backs of the negative kind. Secondly, cybernetics makes the hypothesis that the negative feed-back mechanism explains ' purposive ' and ' adaptive ' behaviour. Broadly speaking what the cybernetic model does for our outlook is to make us understand how purposive behaviour can be manifested by a machine, for " purposive " can now be defined in terms of negative feed-back. An anti-aircraft gun is ' purposive ' in the sense that it is able by means of radar to follow its target.[1] Predictive behaviour is also accounted for by the same mechanism, as illustrated by a gun firing not at a point where an aeroplane is but at a point where it will be—just as a cat chasing a mouse extrapolates the mouse's path.

Support for this striking hypothesis, which is of fundamental importance, is to be found both in rather general resemblances between organisms and electronic machines and in resemblances of more special kinds.

The hypothesis seems to be due mainly to Ashby and Wiener, though others had written on some of the resemblances. Ashby introduced it in a paper in 1940,[2] and Wiener who had the idea for some time put it forward independently in a joint paper in 1943.[3] It was Wiener who introduced the word " cybernetics " in his book in 1948[1]; he coined it from the Greek for a helmsman. Other writers, who, as we shall see, contributed to the detail are sometimes credited somewhat vaguely with the idea, and they may have had it but the literature affords no confirmation of this. It would seem to be fair therefore to regard the general approach as the Ashby-Wiener hypothesis (though possibly Wiener's associate, Bigelow, should also be included).[2]

---

[1] Some readers may find it conducive to clarity to distinguish with Northrop between ' purposive ' and ' teleological '. A machine may start towards a goal, but after it deviates it may make no attempt to pursue that goal. It would then be in one sense ' purposive ' but not ' teleological '. To be ' teleological ' it must continue to seek its goal, which requires negative feed-back. (F. S. C. Northrop, " The Neurological and Behavioristic Psychological Basis of the Ordering of Society by Means of Ideas ", *Science*, New York, 1948, **107**, 415.) In ordinary usage, of course, " purposive " has the sense here ascribed to " telelogical ".

[2] W. R. Ashby, " Adaptiveness and Equilibrium ", *Journal of Mental Science*, London, 1940, **86**, 478-83

[3] Arturo Rosenblueth, Norbert Wiener, and Julian Bigelow, " Behavior, Purpose, and Teleology", *Philosophy of Science*, Baltimore, 1943, **10**, 18-24

4 *Application and Confirmation*

4.0 *Adaptation in Machines.* One of the most important general features, common to organisms and electronic machines, is *adaptation.* Ashby sketched an analysis of it in 1940,[3] which he repeated on subsequent occasions and developed with considerable elegance in 1947.[4] He approaches it through a consideration of the equilibrium of dynamical systems.

In mechanics there are several sorts of equilibrium. A ball lying at rest upon a table gives an example of one sort; if it is given a slight push, there is nothing to make it revert to its initial position or to exaggerate its motion away from there. This kind of static equilibrium is *neutral.* A golf ball at the very top of a sloping roof of a railway coach is also in equilibrium, but this is *unstable*, for if the ball is given a slight push it will move even further from its starting point. Neither of these forms of equilibrium is relevant.

A golf ball lying at the bottom of a hemispherical bowl, on the other hand, is in *stable* equilibrium, for if it is displaced it will return. It may be noticed that the ball, on returning, will overshoot its mark, return once more, again overshoot, and so on, gradually hovering nearer to its initial position. A pendulum at rest but free to swing under the influence of gravity affords a similar illustration. At first sight it is with this kind of dynamical stable equilibrium that

---

[1] Norbert Wiener, *Cybernetics*, New York, 1948

[2] It is interesting to find that Wilbur had part of the hypothesis but did not put it to explanatory use. He writes of a certain mental mechanism: 'The whole system was something on the order of what the radio engineers call a *self-regenerating* system . . . one is entitled to wonder that he [Freud] appears never to have made use of the *well-known* feature in the structure of the reflex arc, the fact that it *feeds back* into itself.' (G. B. Wilbur, " Some Problems Presented by Freud's Life-Death Instinct Theory ", *The American Imago*, Boston, 1941, **2**, 224 n. My italics.) Why well known?

[3] Ashby, loc. cit.

[4] W. R. Ashby, "The Nervous System as Physical Machine: with Special Reference to the Origin of Adaptive Behaviour ", *Mind*, Edinburgh, N.S., 1947, **56**, 44-59

Ashby appears to be concerned, for he gives the example of the pendulum ; but he follows immediately with the illustration of the thermostat, and writes : 'It is an elementary property of all systems in equilibrium that they react so as to oppose disturbance.'[1]

Now the pendulum and the thermostat are examples of two very different states, and in fact the rest of Ashby's argument shows that he is concerned not with dynamical stable equilibrium but with the state reached by a finely adjusted thermostat. He gives an indication that what he has in mind is the important concept known as " steady motion ", and also called " a steady state " ; but he continues, I think misleadingly, to use the word " equilibrium ". A characteristic example of steady motion is that of a top spinning round a vertical axis—the velocities of the system are constant. Similarly, in suitable circumstances, a thermostat may attain such a state, i.e. one in which the temperature remains constant. But a top may spin without its axis being vertical ; its axis will then revolve and describe a cone with its apex on the ground. Then the motion oscillates round the state of steady motion ; likewise, with the thermostat, conditions may give rise to small oscillations round the steady state of approximately constant temperature. It is this concept, of approximation to a state of steady motion, and not the concept of dynamical stable equilibrium, that is relevant. One point here may cause confusion unless stressed. The numerical value of a state of steady motion differs from a constant in being the mean value of oscillations so small as to be approximately constant. The motion involved is thus not constant but smooth.

The physical concept of steady motion may be used to characterise the *state* of adaptation (not the *process* of adaptation) in an organism. The link between them was first suggested by Cannon in 1926[2] and then by Hill in 1930.[3] Their point was that the organism, or a system in it, merely gives the appearance of being in dynamical stable equilibrium, while in fact there is continuous interchange of material and energy which leaves characteristic quantities unaltered. Blood temperature would be an example. Ashby's aim was to give precision to some of the chief concepts involved, notably to develop the characteristics that define " a machine " ; and he gives a very interesting treatment of this.

---

[1] Id., p. 47

[2] W. B. Cannon, " Some General Features of Endocrine Influence on Metabolism ", *American Journal of the Medical Sciences*, Philadelphia, 1926, **171**, 1-20. In this paper, Cannon anticipated several later developments.

[3] A. V. Hill, " Membrane-Phenomena in Living Matter : Equilibrium or Steady State ", *Trans. Faraday Soc.*, London, 1930, **26**, 667-73

He considers a weight suspended by means of an elastic thread. When the weight is given a vertical motion, the weight and thread together constitute a dynamical system. Again if the thread is stretched far enough to break, the falling weight is also a dynamical system. Now ordinarily we should say that we had to do with *two* systems [1] ; but Ashby points out that we can treat them as one system that suffered a great change. He shows that this is easy to justify mathematically. We are familiar with the Newtonian equations of motion for the weight oscillating on the thread ; these may be said to ' describe ' the motion. We are also familiar with the equations of motion of a falling weight ; these may be said to ' describe ' its motion. Now the two ' descriptions ' can be embraced under one wider ' description ' (set of equations) by means of a mathematical step-function. The step-function is a device by which the general ' description ', when a certain variable is below a certain limit, becomes equivalent to the first set of equations, i.e. to the ' description ' of the oscillating motion, and, when the variable reaches that limit, becomes equivalent to the second set of equations, i.e. to the ' description ' of the falling weight. Now a machine in Ashby's sense is such that each configuration of it determines the one following it ; and dynamical systems of the kind discussed are machines in this sense. Thus before the break there was one machine, after it another, and yet the two machines can be regarded as different forms of one machine. We may say that the machine broke and thereby changed from being machine_1 to machine_2, or we may say that the machine_1 broke and became machine_2.

Ashby makes considerable use of this concept of a break in a machine. What he is concerned to show is that such dynamical systems must either have a break or attain a state of steady motion (which, of course, includes dynamical stable equilibrium as a special case). The point of his interesting argument is perhaps intuitively evident—a weight suspended from an elastic thread, for instance, will eventually come to rest unless the thread breaks. Otherwise expressed, the point is that no form of unsteady motion, such as a large oscillation, can persist without the occurrence of a break. What Ashby is doing, however, is not just to show this, but to show that ' breaking or steady motion ' is characteristic of machines in a certain sense. To put his point another way, it may be possible to interpret the concept of a machine in several ways, but Ashby wishes to characterise it so as to answer to a very general model of an important kind.

---

[1] I suppose this is due to a liking for continuity between states of a system.

It is now possible to form a hypothesis about adaptation in organisms, i.e. to consider defining " a *state* of adaptation " in terms of a state of steady motion. That is to say organisms in a state of adaptation would be regarded as approximately in a state of steady motion, like a thermostat. The special case would arise in which the organism would die ; we might then say it had attained a state of dynamical stable equilibrium ; for any effort, such as artificial respiration, would be opposed by other forces until equilibrium was restored. To be contrasted with the *state* of adaptation is the *process* of adaptation ; this would be characterised by the tendency of the system, due to negative feed-back, to resume steady motion.

To make this approach reasonable, it is necessary to enquire whether organisms have precisely the same characterisation of ' breaking or steady motion ', so that they can be regarded as machines in a general sense. There therefore arises the task of finding the characteristics of a machine in organisms. Ashby notes that the impulses at one point affect those at other points, so that the nervous system is a dynamical system.[1] He suggests that breaks may occur in the fineness of certain membranes and the like. He thinks that the excitation intensity of cells is cumulative, so that at a certain point a break would occur. Greater neurological confirmation is needed, but no doubt this aspect will receive further study.

It is doubtful whether Ashby's analogy is wholly correct, for his concept of a machine, broad though it is, is not so broad as it could be. His concept is such that each configuration determines its successor ; but human machines may not always behave in this way. More recently MacKay [2] has shown the possibility of building a machine with an " uncertainty response " ; this important idea probably answers more truly to the human machine, and interesting developments of it may be expected.

It is important to point out that the mechanism of negative feed-back, though it has only recently come to be discussed in physics, is in no way inconsistent with the laws of physics ; on the contrary, it could from the time of Newton have been treated as part of physics. It simply embodies old physical laws in a new way. Thus a machine with negative feed-back, though a new kind of machine, is still a machine. This point will be referred to again in the final section.

---

[1] Ashby, op. cit., p. 57

[2] R. M. MacKay, Communication to the Philosophy of Science Group, London, Nov. 13, 1950. (To be published in a later issue of this *Journal.*)

4.1 *Feed-Back in the Nervous System.* Wiener handled the cybernetic hypothesis in the usual scientific way. He deduced a consequence from it and sought a test. When a thermostat corrects an error it makes a new one, but a smaller one, and this too it corrects by making a further, but smaller error : the negative feed-back leads to closer and closer focus on the target. In certain conditions, however, it will lead to *oscillation.* This comes about if the self-correcting mechanism overshoots its mark too far, i.e. if the error it makes is commensurate with the one it is correcting. Wiener considered that if the organism has negative feed-back this deduction would hold, and that it should be possible to find an organic example of oscillation. His collaborator, Rosenblueth, confirmed this :[1] a patient with cerebellar disease cannot bring a glass of water smoothly to his mouth ; he swings it past on one side, tries to reverse the motion but succeeds in swinging it past on the other side.

> This test is typical of the disorderly motor performance of patients with cerebellar disease. The analogy with the behavior of a machine with undamped feed-back is so vivid that we venture to suggest that the main function of the cerebellum is the control of the feed-back nervous mechanisms involved in purposeful motor activity.[2]

Non-oscillatory motion, of course, characterises ordinary controlled action, such as picking up a pen—' our motion is regulated by some measure of the amount by which it has not yet been accomplished.'[3]

Wiener and Rosenblueth have here found highly important confirmation of the hypothesis.

Further support comes from experiments in which animals have been deprived of a part of the nervous system. Numerous examples are cited by Brazier[4] ; she mentions, for instance, Young's work with cuttlefish in which removal of the higher centres led to unceasing swimming continuing till the fish died ; she also mentions cauterisation of the interpeduncular nucleus in cats, carried out by Bailey, as a result of which they walk forward continuously till they die. Apparently the negative feed-back mechanisms were put out of action or removed.

---

[1] Wiener, op. cit., p. 5

[2] Rosenblueth, Wiener, and Bigelow, op. cit., p. 20

[3] Wiener, op. cit., p. 116

[4] Mary A. B. Brazier, " Neural Nets and the Integration of Behaviour ", in *Perspectives in Neuropsychiatry*, ed. Derek Richter, London, 1950, p. 36

I have distinguished carefully between different types of feed-back, because it might be fruitful to explore the possibilities of each separately. This has led Thompson[1] to look for physiological examples. He makes the interesting suggestion that a positive amplification feed-back is to be found in some cases of renal disease, in which the renal damage leads to hypertension, which in turn causes further renal disease and still further hypertension—a vicious spiral which, once established, proves fatal.

4.2 *Autonomic Reorganisation in a Machine.* Can a machine reorganise itself when faced with a difficulty? Ashby has given a striking design for a machine with this capacity, which has actually been constructed.[2] It is called " a homeostat ".

He points out that negative feed-back is not sufficient to describe all facets of human behaviour. It does not by itself cover learning by experience—the feed-backs have to improve with the passage of time. The corresponding problem for the brain is that as a machine it has to work out an essential part of its own wiring.[3]

Ashby describes the homeostat in some detail, but all that is needed here is to mention some of its properties. Four magnets are suspended in a field due to electric current, so that they are free to move when the current is turned on ; but feed-back mechanisms are present which are operated by the movements of the magnets, and these feed-backs alter the current. This causes fresh movements of the magnets and so on.

The magnets show a definite pattern of movement. In a stable or steady arrangement the four magnets move to a position where they resist attempts to displace them ; if they are displaced, they return. In an unstable arrangement they diverge as far as possible.

Now instead of setting and altering the feed-backs by hand, uniselectors are introduced to do this. They are mechanisms working off a relay current, their function being to alter the way in which a feed-back is set (for instance they might alter the distribution of weight in the pendulum of a clock) ; this can be arranged to happen only when the current from the mains increases to a certain limit at which the relay current is cut off. But the limit is so chosen that the current reaches it only when the system is unstable and the magnets diverge. Hence, so long as instability remains, the uniselectors alter the feed-backs until eventually a stable arrangement comes about. If we knock out some of the uniselectors and alter the feed-backs by hand, the remaining uniselectors will find a stable arrangement. Ashby and his colleagues even played several tricks on the homeostat. For instance they crossed the wires from the mains, whereupon the magnets again achieved stability. In short *the homeostat adapted itself to new conditions.*

---

[1] Alan Thompson, " Cybernetics and Animal Integration ", *The Irish Journal of Medical Science*, Dublin (in the press).

[2] W. R. Ashby, " Design for a Brain ", *Electronic Engineering*, London, 1948, **20**, 379-83

[3] Id., p. 380

Ashby also makes the interesting observation that, while the ordinary electronic machine can perform only those operations for which it was designed, the homeostat holds out the very real possibility of our making a machine capable of performing operations beyond those for which it was designed—and beyond the knowledge of its designer.[1]

At this point it is worth drawing a distinction in connection with those processes that revert towards steady motion after disturbance. Some revert towards the *same* state of steady motion as they were in before being disturbed ; but others are, as it were, content to aim at *any* steady motion. The one is illustrated by the steadiness of blood temperature ; the other by bodily adjustments to a new diet, and by Ashby's machine with breaks :

> A simple breaking down of a completed machine is sufficient, by itself, to produce the objective features of adaptation by trial and error, at least in the simpler cases ; i.e. every machine breaking down automatically re-adjusts its own internal organisation until adaptation is achieved.[2]

His homeostat affords a concrete example. The homeostat embodies both classes of process or tendencies ; the question therefore arises whether it is in line with the use of the word " homeostasis " which was introduced by Cannon.[3] While Cannon was clearly concerned in the main with processes leading to the *same* state of steady motion, like steadiness of blood temperature, he does not seem to exclude from this heading processes leading to *alternative* states of steady motion ; and this seems to be in the spirit of his work. The name " homeostat " therefore seems to have been suitably chosen, as being correlative to Cannon's use of " homeostasis ".

---

[1] This would refute Descartes's principle that there must be at least as much reality and perfection in the cause as in the effect, a principle on which he based one of his proofs of the existence of God.

[2] W. R. Ashby, " The Physical Origin of Adaptation by Trial and Error ", *The Journal of General Psychology*, Provincetown, 1945, 32, 24

[3] W. B. Cannon, *The Wisdom of the Body*, London, 1932

4.3 *Electroneural Analogies.* Neurons are the cell-units of the nervous system, and a set of neurons can form an electrical system. A neuron has a different electric potential inside and outside the cell-wall, and, when these potentials equalise, the neuron is said to " fire ". Impulses are transmitted from one neuron to another through their synapse (that is, the region between transmitting and receiving ' aerials '—an impulse is transmitted by the axon of one neuron and received by the dendrites of a neighbouring neuron). Signals are thus passed in the form of electroneural impulses, and a chain of neurons forms an electroneural circuit. It may be noted that in electronic machines an impulse is either received or not received, a ' gate ' to a circuit is either open or closed, i.e. they work according to the ' all or none ' principle ; and the same is true of the nervous system—a neuron fires or does not fire.

Apparently the brain has alternative circuits available for the passage of a given signal, in the same way as an automatic telephone exchange provides alternative routes for the impulses received when we dial a number. Thus many of the functions of the brain are non-localised, a feature that is paralleled in electronic machines.

Wiener puts forward interesting analogies, which are speculative in that they are not yet confirmed, though there is no reason to suppose that confirmation will fail to come. He thinks that the brain may become overloaded like a busy telephone exchange. The telephone engineer may overcome this by cutting some subscribers' lines or by removing their instruments : the analogue would be the operations of pre-frontal leucotomy and lobotomy respectively. Support for this idea, Wiener mentions,[1] is to be found in the very high level of performance of telephone exchanges that is maintained right up to the moment when overloading causes widespread dislocation ; and the same is often true of mental illness—a patient may show no sign of onset until the last straw makes him neurotic or psychotic very suddenly. Thus Wiener's analogy concerns traffic jams in both the exchange and the brain—too great density of impulses directed towards a certain route or too few alternative routes available. I would add that we could think of the transference in psycho-analysis as a means of earthing blocked charges.

[1] Wiener, op. cit., pp. 175-6

These ideas apply more specifically to memory. Memory does not appear to be connected with any one location in the brain, for widespread destruction of very considerable areas, a frontal lobe or a cerebral hemisphere for instance, does not destroy, though it may impair, memory—that is, previous experiences can be remembered, though sometimes with greater difficulty. It is not settled whether memory is a function of the brain as a whole or is associated with areas in which the duties that should be performed by a damaged part are taken over by the rest.[1] Whichever is the truth of the matter, we may regard memory as 'non-localised' in the sense that damage to a part does not in general wipe out any memories. Storage in an electronic machine is likewise non-localised, but in a narrower sense, namely that alternative accommodation for information is available. This would suggest that in the brain a neuron pool is available, out of which circuits are selected, just as alternative circuits are available at a telephone exchange. And it suggests that a memory requires the requisitioning of a part of the neuron pool, which so long as the memory is not dissipated, is not available for other use. One of Wiener's ideas seems to be that a set of neurons might form a condenser maintaining a charge.[2] Another possibility is that there might be closed circuits of neurons with electroneural impulses travelling round and round, like an oscillation continuing endlessly in an electronic tube. Here we have the conception of what was called "reverberating" chains of neurons by Ranson and Hinsey,[3] following the first hint given by Forbes[4] when he spoke of "a sort of central reverberation" to account for the after-discharge characteristics of reflex contractions; the phrase "self-re-exciting" chains of neurons was used by Lorente de Nó,[5] working independently. A chain is reverberating or self-re-exciting if in a closed circuit one neuron fires, thereby leading its neighbour and in turn the next neighbour to fire, and so on, until the last one makes the first one fire again. In this way activity in the circuit continues indefinitely or regenerates itself, and the circuit thus comes to be called a "regenerative loop".

---

[1] See E. D. Adrian, *The Physical Background of Perception*, Oxford, 1947, pp. 89–92

[2] Wiener, op. cit., p. 146

[3] S. W. Ranson and J. C. Hinsey, "Reflexes in Hind Limbs of Cats after Transection of Spinal Cord at Various Levels", *The American Journal of Physiology*, Baltimore, 1930, **94**, 490–1

[4] Alexander Forbes, "The Mechanism of Reaction", in *The Foundations of Experimental Psychology*, Worcester (Mass.), 1929, Ch. III, p. 152

[5] R. Lorente de Nó, "La Corteza Cerebral del Raton", *Trabajos del laboratoris de investigaciones biológicas de la Universidad de Madrid*, 1922, **20**, 41-47 (not seen by me; referred to by the author in a later paper)

There are two obvious difficulties about these models. If part of the neuron pool is requisitioned, it could not be utilised for other purposes until discharged; thus repressed experiences might have their counterpart in badly blocked regenerative loops. On the other hand, the mechanism cannot be such that when once it is tapped it is completely discharged—else we could recall an experience but once. Thus a regenerative loop would need in addition to have the property of a magnetic tape, which is used in recording machines, in order that the record of stored information should be 'playable' more than once. It remains to be seen whether such a model could be constructed. All these models, moreover, would imply the localisation of memory, and the problem would arise of explaining how memory could survive damage to a tape or regenerative loop.

4.4 *Recognition of Universals.* A particularly important application of these ideas is the attempt to explain recognition of universals, due to McCulloch and his collaborator Pitts. The problem is to know how it is that a square presented to vision at different distances and angles is recognisably the same square.

The intellectualist school of philosophers regarded universals as 'innate' or due to 'natural light'. The empiricists tried to build them up from particulars and naturally failed to produce a reasonable construction. But, if a machine can be made that will 'recognise' forms, we are on the road to a proper empirical account. What is needed is some way of recording forms, no matter how their particular presentations vary.

Two hypotheses about the process have been advanced by McCulloch, one to do with *feed-back reflex mechanisms*, the other to do with *scansion.*

The idea underlying the first is this. If a number of known weights are lying on a table we can calculate their centre of gravity. If, instead, the weights are replaced by a thick sheet of heavy material, whose thickness varies continuously—like a relief model of the countryside—we can still calculate the centre of gravity (by the mathematical method of double integration). Now if the weight of every minute part of the sheet is replaced by a corresponding degree of brightness, the same mathematical method will give the centre of gravity of distribution of brightness. And this is done in relation to some reference point whose mathematical name is "the origin". Now suppose the eye is focused on some point, and calculate the centre of brightness in relation to this point as origin. The result gives the position of the centre of brightness in relation to the fixation point of the eye, and this position is expressed by two numbers (co-ordinates). McCulloch's hypothesis is that the superior colliculus computes this, (i.e. presumably is fitted with negative feed-backs whose function is equivalent to the pencil-and-paper calculation) ; the superior colliculus signals the resulting position, in the form of impulses proportional to the two co-ordinates, to the lateral and vertical eye-muscles ; and the effect of these impulses on these muscles is to turn the visual axis from its fixation point to the centre of brightness. This continues until the fixation point coincides with the centre of brightness.[1]

Thus we have a hypothetical radar-like negative feed-back by which the focus of attention gravitates towards the centre of brightness. McCulloch then concludes : 'If a square should appear anywhere in the field, the eyes turn until it is centered, and what they see is the same, whatever the initial position of the square.'[2] But unfortunately he does not explain how centering a square of brightness is equivalent to 'recognition' of squareness.

Let us now consider the scansion hypothesis. This idea, which was first suggested by Craik,[3] is that the scansion mechanism in television might play a rôle in perception. The first thing to be done is to translate impulses stimulated by visual forms in space into impulses that are successive. The idea is conveyed readily by the metaphor of morse code, by which a visual 'V' is replaced by four impulses along a time-scale : short-short-short-long. In the second

[1] Walter Pitts and W. S. McCulloch, "How We Know Universals : the Perception of Auditory and Visual Forms", *Bulletin of Mathematical Biophysics*, Chicago, 1947, 9, 127-47

[2] The full quotation expounding the hypothesis is as follows : 'We propose that the superior colliculus computes by double integration the lateral and vertical coordinates of the "center of gravity of the distribution of brightness" referred to the point of fixation as origin, and supplies impulses at a rate proportional to these coordinates to the lateral and vertical eye-muscles in such a way that these then turn the visual axis toward the center of gravity. As the center of gravity approaches the origin, its ordinate and abscissa diminish, slowing the eyes and finally stopping them when the visual axes point at the "center of brightness". This provides invariants of translation. If a square should appear anywhere in the field, the eyes turn until it is centered, and what they see is the same, whatever the initial position of the square.'

[3] K. J. W. Craik, *The Nature of Explanation*, Cambridge, 1943, Ch. VI

place, these impulses must be conceived as travelling round a regenerative loop ; for, as McCulloch puts it, ' a train of impulses in a regenerative loop preserves the form of the fact without reference to the one particular moment when it was experienced.'[1]

The scansion process may be briefly described as follows. A screen, consisting of a glass plate with a metallic lining that becomes luminous when struck by electrons, is swept with a jet of electrons, just as one might sweep a window pane with a jet of water from a hose. The screen is part of the bounding surface of a tube, a cathode ray tube, shaped something like an ink-bottle ; we can imagine the screen as the base of the bottle and the stream of electrons entering through the neck. This stream is prevented from being just a diffuse shower by a magnetic field which constrains the stream into a thin jet. Another magnetic field moves this jet to and fro, as one's hand may move a jet of water from a hose backwards and forwards across a window pane. The jet of electrons begins at the top left-hand corner, moves at very high speed in a straight line to a point very slightly below the top right-hand corner, is cut off, begins again just below where it first began on the left, traces another line to the right, and eventually works its way down to the bottom ; it has then completed one ' frame ', and it may complete 25 or 50 frames each second. Next light from an object is focused on the screen, and the density distribution of the photons striking the screen will vary with the light and shade of the object. The interaction between the electron jet and the cluster of photons in every small area where the jet strikes the screen produces an electric impulse which is conveyed from the metallic lining of the screen by a wire. Thus a frame is converted into a stream of varying impulses ; that is to say every small area of the picture on the screen, of whatever degree of light or shade, is represented by an impulse.

The scansion hypothesis now takes the form that the retina functions as a television screen.

If we consider two square objects at different distances from the eye, the same sort of electric impulses will arise from both. But the *time-relations* between the impulses received in perceiving one square will be different from those involved in perceiving the other square—like the difference between a tune played fast and the same tune played slow. Although, therefore, there is an invariant relation between the two sets of impulses—like the invariance of pitch in a tune played at different tempi—there is also this important difference of tempo to be dealt with. And we may express the effects obtained by scansion by saying that different spatial arrangements having the same form are converted into temporal arrangements of impulses with different ' tempi ' but invariant ' pitch '.

---

[1] Quoted by Northrop, loc. cit., p. 413

Even this result is sufficient for some purposes. Thus McCulloch and Pitts have made it possible to listen to the printed page. It was easy to reproduce letters of a given size by sound ; the difficulty was to reproduce letters of different sizes by the same sound ; but this they succeeded in doing.[1]

But this is not sufficient for the theoretical explanation of ' recognition ' of universals. We will suppose that scansion of different squares produces impulses with different ' tempi ' but invariant ' pitch ', and that these two sets of impulses are stored in two regenerative loops. What happens then is not clear. We must suppose either that the two loops synchronise despite the difference of ' tempo ' or else that difference of tempo disappears in regenerative loops. It is hard to believe that the latter is possible. Apart from the difficulty of guessing how it could happen, there exists the recognition of literal invariance of pitch in a tune played at different tempi *where the difference of tempo is also recognised.*

One is therefore driven to suppose that two regenerative loops containing reverberations of invariant ' pitch ' though different ' tempi ' synchronise. But the problem is merely reopened at a new place if we interpret this by saying that the invariant ' pitch ' in different reverberations is in some sense ' recognised '. To proceed further, we might try forming hypotheses connecting similarity of ' pitch ' with our characteristic response of attributing the same form to two different squares.

The hypothesis of scansion is most suggestive, but it cannot be regarded as complete.

4.5 *Deductive inference.* It is well known that logical machines exist. They are capable of giving deductions from data in a logical calculus. There is nothing very striking about this in principle. What is more interesting is the view that electroneural nets of regenerative loops have exactly the same formal properties as the axioms of mathematics and logic.[2] In one sense this tells us nothing new, for after all the brain must be such that we can carry out mathematical and logical inferences. The point of importance is that the concept of electroneural nets should have the formal properties required for this, i.e. some of the important properties it ought to have if it is to provide a reasonably good model of the brain. What this result gives, therefore, is not striking information about electroneural nets but support for the hypothesis of electroneural nets. Unfortunately the details involved here are technical.

---

[1] See Wiener, op. cit., p. 31

[2] W. S. McCulloch and Walter Pitts, " A Logical Calculus of the Ideas Immanent in Nervous Activity ", *Bulletin of Mathematical Biophysics*, Chicago, 1943, **5**, 115–33

In this context we should notice an idea put forward by Craik.[1] He suggested that, when contemplating purposive action, one considers a ' symbolic model ' of the action. Unfortunately his treatment was too sketchy to be illuminating. But he evidently entertained hope of being able to describe abstract conceptions, such as meaning, implication, number, and so on in terms of such symbolic models. There can be little doubt that all this is true in some sense or other ; but whether it is true in the sense he sought, of models that could be built, cannot yet be answered.

5 *The Organism as an Open System*

An interesting and important development that should be brought into relation with cybernetics is the interpretation of the organism as an open system. This is due to von Bertalanffy,[2] who has published a number of writings on system theory, and to Burton[3] who has briefly put forward the same idea.

A dynamical system is ' closed ' if it neither receives nor yields energy (nor exchanges material with its surroundings). For example, a perfect calorimeter containing a needle oscillating to and fro under the influence of a magnet would be a closed system. But the moving needle itself would constitute an ' open ' system, for its energy would be passed to something outside it, say to the surrounding air. The chief link to be established is that between the open system and steady motion. Since an open system is one that exchanges energy with its surroundings, there are the possibilities that (i) it might lose energy until it attained dynamically stable equilibrium, (ii) it might absorb energy until it broke, (iii) it might oscillate, and (iv) it might attain steady motion, or at least perform small oscillations round the state of steady motion. Thus an open system, though it does not necessarily do so, can in suitable conditions attain steady motion. Since we associate the organisms with the same field of possibilities, we must regard it as an open system. In his mathematical and hypothetico-deductive development of the theory of systems, von Bertalanffy deduces that the properties of steady motion characterise in suitable conditions open systems and that they never characterise closed systems. Then, on the hypothesis that the organism is an open system, it must be capable of assuming steady motion ; but we have identified a state of adaptation with steady motion ; we therefore have confirmation that the organism is an open system. What this does for us is not to give us new information, but to give a theoretical framework into which fundamental properties of the organism fit—and there is always the possibility of gaining further knowledge from such a theoretical framework. Thus von Bertalanffy goes on to develop from it the concept of ' equifinality ' for final states, such as size attained by growth, which are by definition independent of initial conditions—he shows, for instance, that a man's size depends mainly on the relation between the weight and the surface area of the body, and not on conditions obtaining in the man's early history. In this way end-results of growth are covered by the same theory of systems that covers the concept of the organism as capable of steady motion.

---

[1] Craik, op. cit., Ch. V

[2] See Ludwig von Bertalanffy, " The Theory of Open Systems in Physics and Biology ", *Science*, New York, 1950, **111**, 23-9 ; and " An Outline of General System Theory ", *The British Journal for the Philosophy of Science*, Edinburgh, 1950, **1**, 134-65

[3] A. C. Burton, " The Properties of the Steady Rate Compared to those of Equilibrium as Shown in Characteristic Biological Behavior ", *Journal of Cellular and Comparative Physiology*, Philadelphia, 1939, **14**, 327-49

System theory is not to be regarded as a part of cybernetics, because it is a very general theory with all kinds of applications. The connection between open systems, steady states, and adaptation brings the subject into relation with cybernetics ; but, in the main, system theory works on its own. Thus it leads to the hypothesis that the second law of thermodynamics does not hold for the organism ; but such developments, interesting as they are, are not connected with cybernetics.

6 *Electro-encephalography*

Has electro-encephalography [1] or EEG any bearing on cybernetics ?

Broadly speaking, what happens is that if electrodes are put to the head, various rhythms or waves, which emanate from various parts of the brain, are recorded by EEG machines in the form of wave-like curves on paper. EEG is used as an instrument of research in psychiatry. Thus an epileptic patient, whose rhythm in a resting state was first obtained, was confronted with a light flashing at regular intervals ; when the frequency of the flashes became equal to or a multiple of his rhythm when resting, his new rhythm showed the characteristic form associated with an epileptic attack. 'This observation appears to support the hypothesis, based on study of analysed epileptic records, that certain types of seizure are due to exact synchronization of cerebral rhythms previously slightly out of step.'[1]

---

[1] For this new subject see *Electroencephalography*, ed. by Denis Hill and Geoffrey Parr, London, 1950 ; see also the journal, *EEG and Clinical Neurophysiology*, Montreal.

Different characteristic rhythms are labelled by Greek letters. Here we confine ourselves to the alpha-rhythm. This is recorded when a subject is relaxed with his eyes closed, and the rhythm alters when the eyes are opened. In general there is nothing surprising in finding such rhythms ; for from electrodynamics we should expect to find waves associated with electric circuits. What the precise connections between these waves and the various electroneural circuits are is unknown. But one particular link is widely assumed, and this relates EEG to cybernetics : it is conjectured that the alpha-rhythm is associated with the scansion mechanism.

Probably the most valuable theoretical use of EEG will be the more exact location of the electroneural circuits in the brain that are associated with specific bodily functions ; but the field of possibilities is wide.

7 *Evaluation of the Cybernetic Model*

We have seen that the cybernetic hypothesis, that negative feed-back mechanisms are fundamental in the nervous system, can be supported in numerous ways. The organism can be shown to have some of the basic characteristics of a machine with negative feed-back. Dislocation of feed-back mechanisms can lead to oscillations such as are found in cerebellar disease. The same behaviour is to be observed in animals when deprived of an important controlling centre of their nervous system as is to be found in a machine when a negative feed-back is removed. A homeostat has been built, which is able to reorganise itself when interfered with. Certain phenomena of both brains and electronic machines are non-localised. The same drastic remedies are applied to both with some degree of success when they are overloaded. Reverberating chains of neurons or regenerative loops act as storage chambers, as do electronic tubes and condensers. Recognition of universals can be at least illuminated, though not fully explained, by the scansion process in television. Electroneural nets of regenerative loops may well have the formal properties required for deductive inference. Organisms may be interpreted as ‘ open ’ systems.

---

[1] W. G. Walter, V. J. Dovey, and H. Shipton, " Analysis of the Electrical Response of the Human Cortex to Photic Stimulation ", *Nature*, London, 1946, 158, 541

The emphasis has been wholly on *negative* feed-back. Nonetheless it is desirable to define “ cybernetics ” as the science of feed-back in the animal, without restricting this to feed-back of the negative kind. Neurology may now take a greater stride forward than any since the basic work of Pavlov. But it may be found that *positive* feed-back plays an unsuspected rôle ; certainly the possibility should not be overlooked. Moreover, feed-back as discussed in cybernetic literature seems to be solely *simple* feed-back. The possibility should also be explored of finding amplification feed-backs, both positive and negative, in which local sources of energy are tapped. It should be stressed, moreover, that while cybernetics may appear to be the basis of neurology, in fact it should be regarded as a branch, though possibly the most important branch, of that subject ; for there are some important mechanisms that do not involve feed-back.

Cybernetics does not, of course, show that human beings are machines. Many thinkers would hold that they are ; but the important question concerns the kind of machine they are likely to be. When Ashby shows that an organism has certain properties of a machine, this does not establish that they have only those properties. And it is obvious to anyone who has ever met a dog or a man that organisms are vastly different machines from electronic calculators.

What cybernetics does is to show us at least one new kind of machine. One way of putting this is to say that they are not machines in the Newtonian sense. Nor, however, are they machines in the Maxwell or in the Einstein sense. A Newtonian machine, for example, is one for which we can calculate velocities and positions resulting from given velocities and positions. Now a cybernetic machine does not violate the principles of such a machine ; it is simply that feed-back mechanisms go beyond them.

From the time of Newton or Laplace the idea was in the air that human beings were machines ; yet this was obviously incongruous. Cybernetics shows why : the Newtonian machine was not an adequate model of a human being. The question now arises whether the cybernetic model is adequate. Clearly it is not ; there are too many ways in which any machine that can be envisaged at present fails to answer to human behaviour. The possibility remains open that with a sufficient number of highly complex feed-backs or with MacKay's mechanism it would turn out to be a satisfactory model. This would require lengthy discussion. On the other hand it is possible that yet another mechanism (or indeed many others), as yet unknown, may have to be built into our model before it can be human or even animal-like. Nonetheless, whatever mechanisms are discovered and built in and however much they pave the way to new *species* of machines, yet these new things will be *machines*.

All that we know at present is that brains are less like levers and gears than like radar and thermostats ; the resemblance even to these is not very close, but it will have to serve until we find something better.

A last point, of great importance, not unconnected with the preceding remarks, is this. It will doubtless seem to many that cybernetics explains purposive behaviour in such a way as to leave the mental aspect of action out of account, that it thereby gives the *coup de grâce* to the mind, and bears out the behaviourist and physiological approach. Now it is true that cybernetics leaves out the mental aspect of action. But the proper inference from this is that cybernetics is in the tradition of physiology and neurology. It does not solve the problem of the relation of mind and body, because it omits one of the parties concerned and suffers from the one-sidedness of its ancestors. That it should prove stimulating and helpful about this problem, however, there is good reason to hope.

J. O. WISDOM

BIBLIOGRAPHICAL NOTE

IT may be useful to comment on the references cited above or consulted. Some are of historical interest only. Some have to do with details of biological mechanisms. But the fundamental papers, which deal with the cybernetic hypothesis itself, are few.

I would put in a plea for simplicity of treatment, especially in this borderland field of enquiry. Wiener's book, for instance, though most fertile in ideas and containing a useful historical survey, has too many defects of presentation (see my review of it in *The International Journal of Psycho-Analysis*, London, 1949, **30**, 133–7 ; and W. R. Ashby's Critical Notice in

Journal of Mental Science, London, 1949, **95**, 716–24). Again the important paper by Pitts and McCulloch (1947) raises the difficulty of dealing with a paper containing technical material from two fields (mathematics and neurology). Though this kind of difficulty is well recognised, it is not always faced. In a borderland subject like cybernetics what is needed is clear and simple accounts, with little taken for granted, even though papers become somewhat long, and even though the reader will feel at times that he is being taught the elements of his own subject.

For the reader's assistance I list the fundamental writings on the cybernetic hypothesis :

1 Arturo Rosenblueth, Norbert Wiener and Julian Bigelow, " Behavior, Purpose, and Teleology ", *Philosophy of Science*, Baltimore, 1943, **10,** 18–24
2 W. S. McCulloch and Walter Pitts, " A Logical Calculus of the Ideas Immanent in Nervous Activity ", *Bulletin of Mathematical Biophysics*, Chicago, 1943, **5,** 115–33
3 Walter Pitts and W. S. McCulloch, " How We Know Universals : the Perception of Auditory and Visual Forms ", *Bulletin of Mathematical Biophysics*, Chicago, 1947, **9,** 127–47
4 W. R. Ashby, " The Nervous System as Physical Machine : with Special Reference to the Origin of Adaptive Behaviour ", *Mind*, Edinburgh, 1947, **56,** 44–59
5 W. R. Ashby, " Design for a Brain ", *Electronic Engineering*, London, 1948, **20,** 379–83

For system theory, see the following, where references to further work on the subject by von Bertalanffy and others will be found :

6 Ludwig von Bertalanffy, " The Theory of Open Systems in Physics and Biology ", *Science*, 1950, **111,** 23–9
7 Ludwig von Bertalanffy, " An Outline of General System Theory ", *The British Journal for the Philosophy of Science*, Edinburgh, 1950, **1,** 134–65

THE NATURE OF SOME OF OUR PHYSICAL CONCEPTS—II*

In this second lecture we shall examine some of the concepts of thermodynamics. We shall be especially concerned to separate the instrumental from the paper-and-pencil component of the operations that define these concepts. It will appear that the paper-and-pencil operations have to satisfy certain demands which in their turn react on the instrumental operations, with the result that in the final picture the instrumental operations have a greater generality than in the original picture.

Let us begin by considering the first law of thermodynamics as conventionally formulated :

$$dE = dQ + dW$$

By choosing to write the law in differential form we have already injected a strong paper-and-pencil component into the situation. This aspect is, however, not of particular importance for our present purposes and has already been sufficiently discussed in the first lecture. We shall be here concerned with other implications. Writing the law in this form implies in the first place that the universe has been divided into two parts : the 'system,' to which the law as written applies, and the rest of the universe, 'external' to the system. 'External' need not be, although it usually is, taken in the sense of geometrically external. In the equation as thus applied, dE is the change of internal energy of the system in any interval of time, dQ is the heat received by the system and dW the mechanical energy or 'work' received by the system from outside, both during the same time interval. If the equation had its ostensible significance, it would mean that dE, dQ, and dW are all to be determined by independent instrumental operations, and that then it will be found as a matter of experiment that the equality written holds. But this is not the actual meaning, for there is no instrumental operation by which dE may be determined which can be formulated so as to apply universally in all situations. The equation as it stands defines dE, it being assumed that dQ and dW are obtainable independently by suitable instrumental operations. The equation, however, is more than a definition of dE; it acquires physical content through the statement appended to the equation that dE so defined is a perfect differential in the variables that fix the state of the system, although dW and dQ taken separately are not such perfect differentials. The concept of 'state' itself is assumed to be already formed, and the variables which determine the state must be known by independent experiment before thermodynamics is applied. What we mean by the 'state' of a system is the complex of all its measurable properties. A state is fixed by its appropriate variables in the sense that whenever the variables resume their former values, the body is again in its former state as determined by all its measurable properties. It is usually understood that the state variables are obtained by measurements made 'now,' that is, that the state variables are not affected by the history of the system. This restriction may, however, be partially lifted in certain special cases, as in a body which is traversing repeatable cycles of hysteresis.

---

* The second of the lectures delivered in the University of London (University College), on 24th, 26th and 28th April, 1950.

From the fact that dE is a perfect differential, the conservation properties of energy, or more properly, change of energy, follow by conventional lines of argument. Application of the equation to any concrete instance can give only the value of the integral of dE from an initial to a final state, that is, a change of energy. Although no general lines of procedure can be laid down by which the change of energy can be obtained in the general case, except through the definition, nevertheless, for every specific concrete system it is possible to give rules by which the change of energy may be found in terms only of the state variables by a suitable combination of instrumental and calculational procedures. These procedures require rehearsal and previous measurement on the specific system, but do not involve an application of thermodynamics. For instance, if the system is an ordinary p-t-v system, then the previous measurements involve a determination of the characteristic equation of the substance and the heat absorbed along some non-isothermal. If the system involves a magnetic field, then a determination of energy will demand measurements of magnetic permeability and of magnetic field strength. No case is known which does not fit into the picture, for it has been possible to devise appropriate procedures to fit every extension to new types of force. This fact has led to a view of the nature of energy which is associated most closely perhaps with the name of Poincaré. According to this view, energy is most appropriately described as a convention, made to suit our convenience, but with no further significance.

Poincaré's argument for this is that if we were ever confronted by a situation in which the conservation of energy apparently failed, we would at once save the situation by inventing a new form of energy. The justification for this point of view is implied in the discussion just given. I believe, however, that this is only a very partial picture and that energy is far more than a mere convention. For after we have invented the new form of energy to meet some specific new situation, it is required that this new form of energy be applicable to all the other situations in which we now find that the new energy appears. For instance, the newly formulated magnetic energy must be applicable in every situation in which a magnetic force appears. But much more than this, there are reciprocal relations involving the new energy and the entire previous thermodynamic universe of discourse before the new energy was recognised. For instance, when magnetic energy is recognised as a new form of energy, we find that thermodynamics demands a relation between the effect of pressure on magnetic permeability and the magnetostrictive change of volume when a substance is exposed to a magnetic field, a relation which is subject to independent experimental confirmation. Taken in its entirety, energy and its conservation are much more than a convention and express a relationship which is in our control only to a trivial extent. And since there is no general instrumental method of measuring or defining energy, but each new kind of situation has to be met by new *ad hoc* invention, it follows that the discovery of each new kind of force constitutes a new crisis from the point of view of the energy concept, the outcome of which is uncertain until the appropriate invention has been discovered. Fortunately, there have been only a small number of such crises in the past, and there is no present indication that we have to fear others, unless indeed we are made uncomfortable by the situation inside the nucleus of the atom.

Instrumentally and physically, energy is defined only through its differential, or, by extension, it is defined only with regard to a path leading from an initial to a final state. Formally, the energy difference between any pair of states will be independent of the path between them. The paper-and-pencil operation of integrating the differential of E along any joining path corresponds exactly to the physical operation of displacing the system along the path. But at this stage, a new purely paper-and-pencil element is injected into the situation. For it is a mathematical theorem that any function of two variables (that is, of the initial and the final state) which satisfies the conditions met here that it is independent of the path, may be analysed into the sum of two functions, one of the initial state only and the other of the final state only. Or formulated mathematically, the solution of the functional equation

$$f(x,y) + f(y,z) = f(x,z) \text{ is } f(x,y) = \phi(x) - \phi(y),$$

where $\phi(x)$ is arbitrary for any one particular x and the form of ϕ is determined in every special case by the form of f. That is, all the experimental values of the change of E for every state couple may be summarised by assigning to every *single* state its appropriate number and taking differences. This use of a single number associated with each single state obviously constitutes an enormous simplification for calculation and no physicist would for a moment entertain the thought of getting rid of the energy function at which we arrive in this way. It must not be forgotten, however, that this use of an energy function of the single state is something to which we have been led by the requirements of our paper-and-pencil operations, not of our instrumental operations, and that by so doing we have interjected something into the situation beyond the scope of our instruments. Almost never is there complete and unique correspondence between the instrumental and the paper-and-pencil domain—here the paper-and-pencil domain contains more than the instrumental domain. If we forget this, we are likely to read back into the instrumental situation implications carried over from the paper-and-pencil situation. This, I think, is what has happened when the energy concept is reified and energy assimilated conceptually to the concept of ' thing.' There can be no doubt, I think, that the energy concept is thus reified in the thinking of many scientists and certainly in the thinking of the man in the street. It is not uncommon to hear such statements as : ' the universe is composed of two things, matter and energy.' This point of view takes particular comfort and support from the Einstein mass-energy relation $E = mc^2$. But strictly, the above equation should be written $\Delta E = c^2 \Delta m$, and the deltas refer to an initial and a final state of a system. Even as applied to the atomic bomb, the mass energy relation by itself cannot tell us what to expect. What the equation does tell is that if a certain nuclear disintegration occurs and if a certain change of mass is found by experiment to accompany this disintegration that then we may expect a certain liberation of energy to accompany the reaction. ' Liberation of energy ' itself implies an initial and a final state. One had no reason to expect that energy could be got out of the atom until it had been established by experiment that some nuclear reactions are accompanied by a loss of mass.

Energy is too complex to be assimilated to a thing and it should not be reified. ' Energy ' is an aspect of the book-keeping by which we keep track of how far we have gone from our original configuration. The physics gets into the situation through the limitations imposed by nature on the sort of departures from the original configurations that are possible, and because of the character of these limitations our book-keeping most fortunately turns out to be gratifyingly simple.

We now consider another implication of the first law in the form $de = dQ + dW$. Since this law is ostensibly of a complete generality it applies not only to the original system but also to any sub-system which can be carved out of the original system in any way, not only by actual physical division, but in the paper-and-pencil domain by drawing imaginary surfaces in any way in the interior of the original system or even in the space outside it. For any conceivable sub-system the equation must have meaning, and this implies that it makes sense to speak of the energy in the form of heat or of work entering the system from outside. ' From outside ' means across the boundaries, and at any particular point of space the boundary may be given any orientation whatever. It would therefore seem that the dQ and the dW must imply the existence of fluxes ; in particular if heat enters a system from outside it must enter by flowing across the boundary, so that there is a flow vector for heat, $\vec{q}$, such that $dQ = -\text{div}\,\vec{q}$, where now we write the dQ for unit time. Similarly, for mechanical energy, there must be a flow vector $\vec{w}$, such that $dW = -\text{div}\,\vec{w}$. The implication is that these fluxes, to which we are thus led by the demands of our paper-and-pencil operations, have instrumental significance. For mechanical energy in its various forms these fluxes are already recognised. If the mechanical energy is ordinary stress-strain energy in an elastic solid, then $\vec{w}$ has a known value in terms of the stresses and the velocities of the elastic solid. If the energy is energy of the electromagnetic field, then its flow vector is given by the conventional Poynting vector. In all cases the flux of generalised mechanical energy at any point in space, either occupied or ' empty,' can thus be determined in terms of specifiable instrumental operations. Furthermore, the operations are made at the point in question. From this point of view the flux of mechanical energy may be said to have physical ' reality '—in fact this is the operational meaning of ' reality ' under these circumstances. Similarly, the flux vector for heat, $\overleftarrow{q}$, may be supposed to have operational instrumental meaning at every point although the situation is more complicated here than for mechanical energy. Inside material bodies by far the major part of heat is transported by conduction. There seems to be no simple instrument that responds to a flow of conduction heat, but such heat flow may be determined in terms of the temperature gradient at the point and the thermal conductivity of the material. In 'empty' space, transport of heat by radiation receives its instrumental meaning in terms of manipulations with perfectly absorbing screens. In ordinary substances, particularly if they are transparent, there must be a certain amount of heat transported by radiation as well as by conduction. I think it is not clear that in all cases it is possible to make a clean-cut separation of the heat thus transported into two parts, or to give a completely satisfactory instrumental specification for determining the total heat transport. In fact, there is a difficulty of principle here, which I think has not been completely resolved, for strictly, temperature is defined only in systems in equilibrium and in such systems heat is transported neither by conduction nor radiation. We shall not attempt to pursue the matter further here, but optimistically hope that the effect is not important in the situations of practice where temperature gradients or radiational fields are not too intense.

We shall accept the implication that the first law as conventionally formulated involves the existence of fluxes of heat and mechanical energy. We see that the existence of these fluxes has been implied by the demands of our paper-and-pencil operations, and that we have then been able to return from the paper-and-pencil domain to the instrumental domain and find there the physical fluxes which did not at first attract our attention. How this should be so we do not inquire—perhaps it is sufficient answer to say that the demands which we exact of our paper-and-pencil operations have themselves a tacit background of long instrumental experience. At any rate, it would appear that paper-and-pencil operations, or mental experiments in general, may have a useful role in suggesting programmes of investigation in the laboratory.

There is an interesting contrast between the relation of the paper-and-pencil operations to the instrumental operations in the case of the fluxes as compared with the case of the energy. We have seen that the mathematics, in introducing a function of single states instead of state pairs for the energy, has introduced a feature which has no counterpart in the physical situation. But with regard to the fluxes the situation is the inverse, because, although the fluxes have full instrumental significance, when it comes to subjecting them to paper-and-pencil operations we discard part of the full instrumental information, and use only the divergence of the fluxes. As far as the mathematics goes, we could add to all the fluxes any constant flux, no matter how large. Here again the mathematics does not exactly correspond to the instrumental situation, but now it is in defect rather than in excess. It is further to be remarked that although conventional thermodynamics utilises only the divergence aspect of the flux of energy in its two forms, a complete logical carrying through of all the verbal or paper-and-pencil implications of the flux aspect of energy eventually leads to the recognition of the tensor nature of energy, as it is treated in relativity theory. Such considerations would, however, take us much further than we can attempt to go here.

Consider now another implication of our paper-and-pencil demand with regard to *dE*, namely, that the first law have validity for every possible method of cutting the system into sub-systems. This implies that *dE* has meaning for every element of volume and this again means that the *dE* must be fully localised. It is not unusual to find the statement that the energy of thermodynamics is not localised, but that all we are concerned with is energy of the system as a whole. While this may be true in any specific case, I do not see how full localisation can be avoided if we are free to sub-divide our systems in any way. The *dE* thus localised may be found by the appropriate instrumental operations within each element of volume. These operations are not general, but have to be found by preliminary rehearsal. In the first instance, at the stage when the rules are being found by which the *dE* in any specific situation is to be determined, the first law is to be regarded as an equation of definition for *dE*. But once the rules have been found in terms of the state variables of the specific substance, the equation when applied in other situations loses its definitional character and acquires physical content. For now the equation states that the *dE* and the *dQ* and the *dW*, all determined by independent instrumental operations, will be found to be connected as the equation requires in all applications.

We have seen that the equation of the first law demands full localisation of energy, or more properly change of energy $\int dE$, and that this localised energy may be determined by instrumental operations. This is all highly satisfactory, but there are other aspects of the situation which are curious and perhaps not in accord with our physical intuition. Consider the localisation of the energy in a system of electrostatic charges and in a system of gravitationally attracting masses. In the electrostatic system the energy is localised in the field in amount $\frac{1}{8\pi} E^2$ in each unit of volume. The energy so localised may be determined instrumentally at each element of volume ~~merely~~ by measuring the E with ~~the~~ proper instrum~~ent~~ and making the calculation. Furthermore, when the configuration of the electrostatic system changes, the energy localised in an element of volume changes, which means that energy flows out of the element of volume. This flow of energy may be detected instrumentally by measuring the magnetic field in addition to the electrostatic field during the change and combining the electric and magnetic vectors in the well-known way to get the Poynting flux of energy. The situation is entirely different in the gravitational case, although mathematically the forces are given by inverse square laws in both cases. For in the gravitational case there is no detectable effect corresponding to the magnetic field, so that there is no flux of energy through empty space when the gravitational configuration changes. The result is that we cannot localise gravitational energy in the space surrounding the masses, but have to localise it in the masses themselves. This we can consistently do. If a gravitating system freely changes its configuration without constraint, then as the masses accelerate, there is a conversion of potential energy into kinetic energy within each mass, the total energy localised within each mass remaining constant. Or if the gravitating masses approach each other under constraints and so under approximately equilibrium conditions, then the necessary transport of gravitational energy from one mass to another takes place through the elastic members which serve as constraints, according to the mechanical Poynting vector in terms of the stresses and velocities. In no case do we have to have flux of gravitational energy through the surrounding 'empty' space, and this is consistent with the empirical finding that there is no instrumental indication of such a flux in space.

The situation is thus far satisfactory. But it is not so satisfactory when one considers that we would have been obliged to give the localisation of the gravitational case to the electrostatic case if it were not for the known magnetic field surrounding charges in motion, and that these magnetic fields are difficult to detect and were demonstrated only comparatively recently by highly skilled experimentation. Hence, at first, one hundred and fifty years ago, electrostatic energy would have had to be localised in the charges ; this is indeed possible mathematically as is demonstrated in any text on electrostatics. However, it does not feel right physically that so drastic a change of localisation of energy, from residing in the charges to being dispersed through infinite space, should be the result of the discovery of the mere existence of a force so small that for a long time it was below the possibility of detection. In fact, the gravitational situation appears to me so repulsive to physical intuition, which I suspect means merely my tacit verbal demands, that it seems to me we may well be on the lookout for a new gravitational effect, analogous to the magnetic field, too small to have yet been detected.

One may also feel uncomfortable in the face of other aspects of the energy situation. It does not feel right that energy should continually circulate indefinitely in closed paths in the complete static case of a superposed electrostatic and magnetostatic field. It is also to be remembered that although energy ‘ flows,’ we can assign no identity to the flowing energy, so that we cannot say that the energy which was there is now here, and therefore we can assign no ‘ velocity ’ to the flow as we can to the flow of a thing. This points up the remark we have already made. Although energy is in certain aspects like a thing, in its conservation and now in its localisation, it is certainly not like a thing in all respects, and is not to be reified.

Another feature in the energy situation which has unsatisfactory aspects, is that the flow vectors, $\vec{q}$ and $\vec{w}$, are not completely determined by the demands that we have put upon them. For all that we have demanded is that they be such as to give the correct dQ and dW, the latter being scalar quantities associated with the elements of volume. This amounts mathematically to the only demand put on the $\vec{q}$ and $\vec{w}$ being that their divergencies should have the proper values. But a vector is not uniquely determined by its divergence. On the other hand, our physical intuition demands that these flow vectors be capable of instrumental detection, and this means incidentally that they be unique. It is true that we have seen that such unique instrumental specification is possible, but to the extent that the uniqueness is not necessary we have not fully employed our instrumental information in constructing our thermodynamics.

The energy defined by the first law has, as we have discussed it thus far, great generality, but it is subject to at least one limitation as shown by the formulation, namely, we have defined it only in circumstances such that dQ and dW have meaning, or what is the equivalent, circumstances for which $\vec{q}$ and $\vec{w}$ have meaning. Now I think it is evident that the total energy flowing into a system cannot always be analysed into a thermal and a mechanical part. If one goes back to the original definitions of calorimetry he will see that the concept of heat and heat flow were formed under the express restriction that there was no motion of the parts of the system. That is, it is assumed that the heat which enters a system is measured in the absence of mechanical effects and similarly that the mechanical energy is measured in the absence of thermal effects. This last restriction is particularly applicable in radiational fields. How shall we separate the two effects when we have both mechanical and thermal effects, as in the convective cooling of a body in a rapidly stirred fluid with small scale turbulence ? With what instrument shall we separate thermal from mechanical effects in a medium with small scale turbulence ? Or, if there is frictional generation of heat on the surface of a system, the analysis into heat and work flow encounters infelicities.

How shall we handle situations in which we cannot assign a definite instrumental meaning to heat and work ? Let us suppose in the first place that we are dealing with systems composed of substances for which the dE has been established by previous experiments for every specific substance as a specific function of the variables of state. This means that for any configuration of the system we are in a position to localise completely the energy as a scalar in every element of volume, and our only source of embarrassment arises because we do not know how to specify adequately the processes by which the localised energy changes. We at first deal with this situation formally and verbally. We lump together the $\vec{q}$ and the $\vec{w}$, unknown in detail, into a generalised flux of energy, which has to satisfy the requirement that its divergence is equal to the time rate of change of the localised energy. This latter requirement is merely the mathematical form of the statement : 'If energy is not created, then the increase of energy in unit time within any element of volume is equal to the amount of energy which has entered the volume across its boundaries in the same time.' This is, I think, in the first place a purely formal verbal demand which is imposed by the necessities of our thinking apparatus. If we assume that the increase of energy has independent instrumental meaning, as now it has, this formal statement contains two terms, otherwise unknown, which are to that extent defined by the statement itself. These two terms are ‘ creation ’ and ‘ energy entering across the boundaries.’ Since we are going to operate under the assumption that energy is conserved and therefore not created, we are left with a single undefined term, namely the flux of energy across the boundaries. Mathematically, a flux can always be found which shall have a specified divergence, so that our method of treatment involves no inconsistency. But the scheme acquires physical significance only when we can find some instrumental operation which independently measures flux of energy across the boundary. I think that a generalised instrumental operation for measuring flux of energy as such, without an analysis into thermal and mechanical components, has not been devised up to the present, and that therefore the device by which we save the flux of energy situation remains on the formal level, where it serves the function of giving consistency and completeness to our paper-and-pencil treatment. This does not mean at all that the other aspect of energy, that is, energy as a scalar localised in the volume elements of the system, is confined to the formal level ; for this aspect of the energy concept is susceptible of the most thoroughgoing instrumental verification. For if we suppose a universe of substances for which their scalar energy functions have been determined in terms of their variables of state, we may then select from this universe every conceivable combination in every conceivable initial condition, imprison any of these combinations in an adiabatic enclosure and check by instrumental measurement that the sum of the energy functions for every stage of all conceivable interactions is constant. In fact, the number of degrees of freedom attending all possible reactions in an adiabatic enclosure is so great that we may invert the process, and by studying the adiabatic reactions of a new unknown substance with other known substances, find how to assign an energy function to the new substance. In this way we could deal with a substance which takes part in no processes for which the flow of heat and mechanical work are separately defined. But always in the back-ground, in order to get started, we have to assume some substance whose energy function has been established by processes for which heat and work have meaning. In practice this imposes no restriction at all.

In practice, when dealing with unfamiliar forms of energy flow, we may invert the sequence suggested by the definition, and having found the scalar energy function for some substance, use that to determine the energy flow. This is essentially what is done when the flow of radiant energy is measured in terms of the rise of temperature of a screen placed in its path. It does not appear, however, why the energy flow determined in this way should be described, as it usually is, as a flow of radiant *heat*. There are situations in which this method of description may lead to infelicities, particularly when the flow is unidirectional. It would seem preferable to describe this flow as, at least partly, a flow of mechanical energy rather than as heat.

It is interesting to observe how the significance of the equation of the first law, $dE = dQ + dW$, for those systems to which it is applicable, has changed as use of the equation has extended. It began as a formal equation of definition of dE, and therefore with a strong paper-and-pencil component of significance. It has ended as an equation in which each term has independent instrumental significance, and therefore as an equation which in any specific situation is capable of complete instrumental verification. As such, it has evolved from a definition to a statement of a ' natural law.' We may, if we like, take this as the definition of ' natural law.'

The question of separating energy in the form of heat from that in the form of work is not of great importance when we are dealing with the first law alone, but it is vital when we deal with the second law. A completely satisfactory method of separating the two concepts of work and heat, even in principle, has not I believe yet been found. At least one serious attempt to place thermodynamics on a logically rigorous foundation by the use of the postulational method, with which I have been familiar, failed because a distinction in principle between heat and work could not be formulated with sufficient precision.

We now turn to a consideration of certain aspects of the second law. We start with the conventional mathematical formulation of the law in the form : $dS = dQ/t$. Here dS is the change of entropy of the system in question for any reversible process during which the heat dQ is absorbed. The equation as it stands is at first a definition for dS. But just as in the case of the first law, the equation is supplemented by the statement that the dS so defined is a perfect differential in the state variables of the system. It follows that the total increment of entropy, obtained by integrating dS, on passing from any fixed initial state to any fixed final state is independent of the details of the process and depends only on the initial and final states. Again, as in the case of the energy, it is convenient mathematically to deal with this situation by choosing some arbitrary state as the initial state and assign an arbitrary value of entropy to it. By application of the second law the other states of the system may then be assigned single entropy numbers and the difference of entropy between any two states may then be found by taking the difference of the corresponding entropy numbers. With this understanding it is convenient to speak of the entropy as a state function, but as with the case of energy, this must not obscure the essential fact that we are strictly concerned with a state couple and that it is only differences of entropy that have physical significance or are uniquely defined.

The entropy may be determined as a function of state for any specific substance for such states as can be reached by reversible operations. The entropy having been thus determined, the equation of the second law changes its character, becoming no longer merely an equation of definition, but it acquires complete physical status and states an equality which will be found experimentally to hold between two independently instrumentally determined things, namely the difference of entropy between any two states and the integral taken over any reversible path connecting them. The necessity for finding at least one reversible method of reaching the final from the initial state does not in ordinary practice lead to any onerous restriction, for such a reversible process can usually be discovered. There are cases, however, where the restriction is onerous, and there are even cases where it seems impossible in principle to remove the restriction. We will return to this question.

Just as in the case of energy, the entropy to which we are thus led must be accorded full localisation in space, because the system can be divided into subsystems in any arbitrary way and the second law applied to every arbitrary element. This demands the existence of fluxes of entropy to account for the changes of localised entropy, just as we had to have fluxes of energy. The necessity for the existence of a flux of entropy is at first a consequence of a purely formal and verbal demand. The formulation of this demand is a little different in the case of entropy than for energy because we do not have conservation of entropy. The demand in the case of entropy is thus more general than it was for energy and may be formulated as follows: ‘The excess of entropy in any given region of space at the end of any interval of time over that which was present at the beginning of the interval, is equal to the entropy which has entered the region across the boundaries in the same interval, plus that which was created within the region during the same interval.' As it stands this is completely formal and verbal and devoid of any factual content, but I think it represents a form of thought without which we could not conveniently operate. It is formal because we have not as yet any instrumental method of giving meaning to flow of entropy across the boundaries or to 'creation of entropy.' We now seek to give factual content to our formal and conventional demand by discovering instrumental operations corresponding to flow and creation. Whatever instrumental operations we discover must be subject to the restriction that their relation to each other is consistent with our formal demand.

It is only comparatively recently that the recognition has been growing that the full thermodynamic picture presented by both our instrumental and paper-and-pencil operations demands the existence of a flow vector for entropy and also a 'creation function.' Not only this, but it is coming to be realised that flow and creation provide the tools by which certain irreversible processes may be brought within the ambit of thermodynamics. It is notorious that classical thermodynamics has only a very limited applicability to irreversible processes. Beyond stating that the total entropy of the universe increases during any irreversible process it can do very little ; it does not attempt to state how great the increase of entropy is or where it is situated. Physics should not be thus impotent, and it does not accept such impotence in other fields. We may recognise two kinds of irreversibility. The first arises from too great complexity of the process, as when a gas expands turbulently through an opening from a region at one pressure to a region at a lower pressure. Here our failure to deal with the situation arises from our inability to describe exhaustively what is going on in terms of measurements with the large scale instruments which measure the variables of state.

In addition to these complicated processes, there are simple irreversible processes which can be completely specified in terms of a few simple measurements and it would seem that we ought to be able to deal with these. Perhaps the simplest of all such irreversible processes is the steady flow of heat by conduction down a temperature gradient. Let us analyse this simple situation and see whether we can find a plausible way of assigning a flow of entropy and a creation of entropy. We may imagine the heat conduction taking place through a conducting bridge connecting two heat reservoirs at different temperatures. The reservoirs themselves are idealised as usual by supposing them made of material which is an infinitely good conductor of heat. This ensures that the heat imparted to the reservoir from outside is uniformly distributed through it by a reversible process. In the final steady state, the hot reservoir is continually losing entropy because it is losing heat, and the cold reservoir is continually gaining entropy. The conducting bridge remains in a steady state, and because of our principle that its entropy is a function of the state variables, its entropy is not changing. If the quantity of heat Q passes by conduction from the hot reservoir at temperature t_2 to the colder reservoir at temperature t_1, the loss of entropy of the hot reservoir is Q/t_2 and the gain of entropy of the cold reservoir Q/t_1, and since the conducting bridge does not change, the gain of entropy of the universe is $Q(1/t_1 - 1/t_2)$. We must say that the entropy of the cold reservoir increases because entropy flows into it, and the entropy of the hot reservoir decreases because entropy flows out of it. By the principle of sufficient reason we must say that entropy flows into the reservoir where the conducting bridge joins it. If $\vec{q}$ is the flow vector corresponding to Q, that is, the vector given by Q divided by the time of flow, then these demands force us to say that the entropy flow vector where the conducting bridge connects with the colder reservoir is $\vec{q}/t_1$, and similarly the flow vector where the bridge connects with the hot reservoir is $\vec{q}/t_2$. We would be going out of our way not to suppose that in general, whenever we have heat flow by conduction, the corresponding entropy flow is $\vec{q}/t$. Apply this vector now to the conducting bridge and consider an element of volume bounded by two surfaces perpendicular to the axis of the bridge. The vector $\vec{q}$ is constant all down the bridge. Since the exit temperature of our element of volume is lower than the entrance temperature, our expression for vector flow demands that the entropy flowing out be greater than the entropy flowing in. But since, as we have already seen, no entropy is accumulating within the region because the state is steady, this can only mean according to our formal demand that entropy is being created within the region. On reflection this appears as it should be. For the volume element is the seat of a continuous steady irreversible process leading to the continual entropy increase of the universe and what more natural than that the entropy should be created in the region where the irreversible process occurs? But the entropy having been created, does not remain in its place of origin, but is immediately transported away to find its final resting place elsewhere, in the reservoirs.

It is easy to find a general mathematical expression for the entropy created which reduces to the correct value in the simple case just described. This is :

$$\begin{pmatrix}\text{Entropy created per unit} \\ \text{time per unit volume}\end{pmatrix} = -\frac{\vec{q} \cdot \text{Grad}\,\tau}{\tau^2}.$$

It is natural to postulate that whenever we have conduction of heat, which is always an irreversible process, we have a creation of entropy according to the equation just given, whether or not the conduction process is accompanied by other reversible or irreversible processes. This postulate gives to 'creation of entropy,' in the case of heat conduction, the independent instrumental significance that we have been seeking. If other irreversible processes occur jointly with heat conduction, and if we can similarly find the 'creation' to be associated with them, then we have given full instrumental status to 'creation.' In addition to the creation function, we have also, in the case of thermal conduction, found the flow vector. Let us suppose that we have similarly found an appropriate flow vector for other processes. Then we have given full instrumental status to our statement which in the first instance was formal and conventional, namely : 'Entropy accumulating within a region of space is equal to that created within the region augmented by the amount which has flowed in across the boundaries.' We have now a statement which is capable of complete instrumental verification. Let us assume that the experimental verification is complete. How shall we describe our result ? I think we would not say that we have thereby verified the statement, but rather that we have verified the correctness of our expressions for creation and flow of entropy. The form of the statement is, I think, one to which we would cling no matter what the experimental situation. The form thus has only a verbal and not a physical significance, but there is, nevertheless, a certain compulsion associated with it.

Going back, there are other simple physical processes beside thermal conductivity for which it is possible to find appropriate creation and flow functions for entropy. A particularly simple case is the generation of heat by an electrical current flowing against a resistance. This is an essentially irreversible process, for which it provides that the appropriate expression for the creation of entropy per unit time per unit volume is simply ri^2/τ, where r is the specific resistance of the material. The heat generated by the current flows away from the place of generation by conduction and thus provides an entropy flow vector,

$\vec{q}/\tau$ by which the entropy ri^2/τ is carried away. This entropy flow is adequate to carry away the generated entropy, and thus may be taken as the entropy flow vector, and thus may be taken as the entropy flow vector for irreversible entropy generation by a current in overcoming electrical resistance, except for a small term, which can usually be neglected, to which we shall return in the third lecture.

The creation function for entropy being known for thermal conduction and electrical resistance, we are in a position to check the correctness of our views by application to a special problem. This is the problem of the thermoelectric circuit. There are certain relations between various thermal effects and the electromotive force in such a circuit which are known empirically and are so simple that we should be able to deduce them. Since the entire source of the electrical energy is thermal, we should be able to deduce the relations by a thermodynamic argument. The relations were thus deduced by Lord Kelvin, but only by an argument which he himself recognised to be lacking in rigour. What Kelvin did was to entirely disregard the irreversible processes in a thermoelectric circuit, and deduce the relation by supposing that only reversible processes occurred. Now this is usually legitimate in the elementary situations of thermodynamics. For usually it is possible to give the system such dimensions or otherwise modify it so that the irreversible aspects may be made vanishingly small in comparison with the reversible ones. This is ordinarily done, for example, by conducting the processes infinitely slowly, or by using infinitely good conducting material for the reservoirs, or some similar device. Now it turns out that the thermoelectric circuit occupies a quite unusual position with regard to the possibility of making the irreversible processes of vanishing importance. For there are two irreversible processes in the thermoelectric circuit, thermal conduction, and development of heat by Joulean heating, which are so connected by the physics of the situation that they cannot both be made vanishingly small simultaneously, but when one is made small by suitable choice of dimensions or otherwise, the other becomes large. It follows that both cannot be made to vanish, so that the total amount of dissipation in the system compared with the useful utilisation always has a finite value. An upper limit can be set rigorously to the ratio of these two effects by the fundamental inequality of classical thermodynamics. But when the numerical values are substituted which correspond to this inequality, one side of the inequality turns out to be one thousand times larger than the other, so that no useful information at all is given by the rigorous expression. This was first shown by Boltzmann.

No such difficulty arises, however, if this problem is treated according to the point of view that each of the two irreversible processes is accompanied by its own characteristic creation of entropy, unmodified by the occurrence of the other processes. For now we can write down the entropy generation of the entire process, including both reversible and irreversible parts, and from this eliminate the contribution arising from the irreversible parts, because now we know its exact quantitative value, and remain with a relation between the reversible parts. This turns out to be exactly the relation derived by Kelvin by unsatisfactory argument, and furthermore is checked by experiment with a precision which becomes continually better with each improvement in the measurements. The same sort of argument can be extended to similar situations, and in my book on *The Thermodynamics of Electrical Phenomena in Metals* will be found several applications, particularly to the deduction of transverse thermoelectric effects in crystals.

The same method of handling irreversibility has essentially been employed by Eckart for two other processes so simple that they can be completely characterised in terms of a few measurements with macroscopic instruments, namely the processes of viscous flow in a fluid and of diffusion down a concentration gradient. In his analysis Eckart has found how to write explicit expressions for the entropy created when the process runs. The expressions are simple and appear intuitively natural. They localise the creation of entropy where it would be expected, that is, where the irreversible process occurs. The permanent abiding place of the increase of entropy is, however, elsewhere, the entropy being conducted away on the flow vector, from its place of origin as soon as created. Eckart also found the appropriate expressions for the flow.

In a recent paper Tolman and Fine summarised these applications and have emphasised the creation and flow aspects of entropy. In Belgium, De Donder and his school, notably Prigogine, have been much occupied with such questions, particularly with irreversible chemical reactions. But as far as I know all such analysis up to the present has been confined to cases where there is some method of assigning an entropy to the material within which the irreversible process occurs, that is, where some reversible method exists for getting from the initial to the final state. If the entropy can thus be independently evaluated, it is a comparatively straightforward problem to work out the creation and flow functions. But I believe that the general problem of irreversibility is not as simple as this and that the fundamental difficulty in the general case is in knowing whether an entropy exists, and, if so, in finding how to evaluate it. For, going back to the original definition, entropy is defined as a state function only for those states for which at least one reversible method exists for getting from the standard state of reference to the state in question. But many of the bodies of experience are in such a state that every change to which the body is subjected contains some component of irreversibility. An example is any piece of metal which has been heavily strained plastically. That is, many of the materials of daily life, perhaps most of them, are completely isolated by irreversibility from the universe of thermodynamics, so that in principle the classical concept of entropy is not applicable to them. Incidentally, the whole living world of biology is thus isolated. Strictly it cannot be asked whether living things violate the second law of thermodynamics, because the entropy concept, in terms of which the second law is formulated, does not apply. For if the entropy concept were applicable, it would involve some reversible method of getting from the non-living to the living, and this involves at the least the artificial creation of life, an operation which we cannot at present perform.

Returning to the plastically strained metal, strictly we cannot speak of its entropy, and if we find how to do so, it must be by an extension of the classical entropy concept. Now it would be most desirable if we could so extend the entropy concept. One reason is that statistical mechanics represents entropy as one of the most universal of the properties of matter and sets up specific methods for calculating it, with no reference to reversibility or irreversibility. It would be desirable if some connection could be found between the statistical concept and the thermodynamic entropy concept in this more general setting.

What requirements shall we place on an extended thermodynamic entropy? It must, of course, reduce to the known values in the limiting case of reversibility. In addition, during the irreversible process it would seem that we must also demand that the total entropy of the universe always increase at every stage of the irreversible process and that the region of entropy creation be the region where the irreversibility is taking place. These demands are perhaps not sufficient to give us a unique solution, so that in a first exploration of the possibilities we will naturally choose a solution which appears most simply to satisfy the conditions. I have examined the possibility of such an extension to several comparatively simple non-steady irreversible processes, namely the cyclic deformation of a body with stress-strain hysteresis, the deformation of the conventional idealised perfectly plastic body satisfying the von Mises criterion, and a body in which an internal irreversible generation of heat may take place as in the unmixing of a mixed crystal. In all these cases it is possible to find a simple expression for the rate of entropy creation and also for the entropy to be associated with the body in which the irreversible process occurs. Whether the precise expressions found by the simple analysis are the correct ones, will have to be found by experiment. At present the experimental basis has not been sufficiently laid for a thoroughgoing critique, and we have to be satisfied with the crude indication that there seems to be no glaring inconsistency.

It would appear, therefore, that there is nothing intrinsically contradictory in attempting to generalise and extend the entropy concept beyond the range of its classical definition and into the range of intrinsically irreversible processes. For each special kind of irreversible process the details will have to be worked out. Entropy has no more a universal instrumental significance than does energy. This detailed working out will involve a background of complete experimental knowledge, which means the ability to exhaustively describe the phenomena with macroscopic measurements. Once having acquired such command, it would appear that we may attack the problem by methods in the spirit, if not the letter, of classical thermodynamics. A final failure of the thermodynamic approach may be anticipated only in those cases where we can not acquire sufficient command of the situation to be able to exhaustively describe it.

P. W. Bridgman

ON THE PRINCIPLE OF ACTION BY CONTACT

I

THE idea of contact-action is an essential part of the theories of classical physics, at least when they are formulated in terms of the field concept as required by relativity. But can or need the principle be reconciled with quantum mechanics ?

The exhaustive description of quantum phenomena reveals so-called *causal anomalies*.[1] The double-slit diffraction experiment performed with electrons is a well-known illustration. If the wave-language is chosen to describe the events on the screen, we must say that the extended wave front is suddenly contracted into points, i.e. the scintillations. Or we may describe what happens at the slits in terms of the particle-language. Then we cannot avoid saying that an electron passing through one slit is influenced by the presence of the second slit since the diffraction pattern (i.e. the distribution of scintillations on the screen) is seen to differ whether one or two slits are used. In other words, we seem to require an instantaneous action at a distance if an exhaustive description of quantum phenomena is given, and this constitutes the causal anomaly.

This has long been known and it is now understood that these anomalies are of a specious sort. If we reject the exhaustive in favour of a restrictive description no reference to causal anomalies need be made. Such a restrictive description speaks only about what is actually measured in a given situation and does not mention how action is transmitted or whatever is not observed under the circumstances. This kind of description can always be carried out since the uncertainty relations in fact rule out the interpolation to unobserved events, e.g. the happenings at the slit when the screen is observed. And we cannot describe the events in terms of two different, complementary, language-systems at the same time. On the screen, we may see at high intensity a diffraction pattern and so we need not refer to the scintillations, i.e. the particles ; while at low intensity we see scintillations, and no reference to the wave need occur. It is always possible to eliminate causal anomalies if we do not mix particle-language and wave-language and if we restrict ourselves to the description of actually observed events. This restriction on the use of classical language avoids violation of the principle of action by contact.

---

[1] H. Reichenbach, *Philosophic Foundations of Quantum Mechanics*, University of California Press, 1944

2

These anomalies are not due to the change from the deterministic to the statistical type of law brought about by the uncertainty relations. The laws of quantum mechanics differ from those of classical physics, but they still furnish a description and prediction of phenomena. This has led to the now customary, technical explanation, or explication, of 'causality.' The requirements of causality are said to be satisfied if we can state a law which allows us, given initial and boundary conditions, to describe and predict a certain class of events, i.e. those events which the theory, and the law, are intended to deal with. The explication is obviously based upon the mathematical structure of physical theory which is usually given by differential equations. This formulation may thus be regarded as specifying a *syntactic definition* of 'causality.'

The analysis of the concept of causality requires a further step. The deterministic conception is given up in quantum mechanics as a result of the uncertainty in our data and the statistical interpretation is introduced. But there still remain assumptions implicit in the use of causal description originating from classical theory that need not apply in other theories. The determinism of classical physics includes at least two assumptions. First, it is assumed that exact data are available ; second, that cause and effect are connected by physical processes occurring as a continuous chain of actions, or by an action by contact.

Einstein pointed out that quantum mechanics as we understand it today is not compatible with the idea of contact-action.[1] This notion has been of great value in classical physics and is a major constituent in the field conception of relativity theory. Physicists generally believe that relativity must apply also to the description of micro-physical events—in the somewhat limited manner in which it has already been used there it has proved most fruitful—and so we seem confronted with a serious difficulty.

---

[1] A. Einstein, in *Dialectica*, 1948, **2**, 307

3

Newton's system of mechanics does not require the concept of *contiguity*. This is necessary only in field theory which involves the finite speed and the continuous character of action. It is not surprising that quantum theory which introduces a discontinuous action through Planck's constant cannot be reconciled easily with the idea of contact-action. It is because this idea is an integral part of classical language that we must restrict the sentences admitted to quantum-mechanical language so as to escape reference to causal anomalies.

The basic concepts of classical and quantum physics differ, and so it is natural that the laws are equally changed. One example is the concept of particle : the behaviour of the Newtonian mass-point differs from that of the elementary particle. What we mean by ' causality ' depends, however, upon the theories to which a causal law belongs. Any discussion of causality in general terms tends to become meaningless, and only specific laws in their context can serve as example. Traditional epistemology attempts to formulate a universal principle of causality governing the whole of nature and so introduces an irrelevant ontology. Causality cannot be regarded as a universal mechanism, and causal explanation does not require to answer the pseudo-question : What is nature really like ? The Aristotelian terminology of the four causes is superseded by the epistemological conception of cause mainly founded on Newtonian mechanics and refined through relativity theory, but modern physics requires a *semantic* conception.

4

Classical physics, particularly field theory, has taken causality to be exemplified by *one basic process*. The causal chain is transmitted from point to point by contact-action, and so the *inverse-square* law becomes the prototype of causation.

By taking a special law as *the* law of causality we commit ourselves to a specific theory and to a definite interpretation, or model. Of course, we must do so whenever we apply the concept of causality to physical processes, or we must specify the law. In application the purely syntactic definition of ' causality ' is supplemented by a *semantic definition* since reference to the physical events is needed. And such definition depends upon the theory, i.e. the language-system, in which a causal law is formulated.

It is well known that the conception of causality as contact-action is closely connected with the ideas of space and time as given by field physics. In relativity theory the time order is determined by the properties of the causal chain. And the relation of the three-dimensionality of space to the inverse-square law was pointed out by Exner many years ago, although this suggestion may be supported by an argument from geometry alone and so is compatible with action at a distance.[1] In other words, the principle of action by contact is essential to the field language in which classical physics is usually expressed.

The classical model of causality arises from the idea of contiguity introduced by field theory. Hume said in the *Treatise* : ' Whatever objects are considered as causes and effects, are contiguous ; and nothing can operate in a time or a place, which is ever so little removed from those of its existence. We may therefore consider the relation of contiguity as essential to that of causation.'

A specific mechanism is introduced for causality. This makes it a physical process ; but it also introduces a special interpretation tied to the classical ideas of space, time, and motion. Is this classical interpretation the only one we can admit to science ?

Again, to quote Hume : ' We have no other notion of cause and effect, but that of certain objects, which have always been conjoined together . . . we cannot penetrate into the reason of the conjunction.' This argument is, of course, mainly directed against the necessitarian interpretation. But it is clear that any statement about causality which asserts more than the regular succession of events goes beyond the strictly logical, or syntactic, definition of ' causality ' and so involves a semantic definition, i.e. a theory or semantic language-system in which the law occurs as a theorem. The inverse-square law when taken as contact-action serves as a model of causality by showing how the causal process works, and so the danger of using ' causality ' tautologically is avoided. But this restricts the discussion to a special theory, in this instance, to classical field theory. If a causal law is to be testable it must be a constituent of a particular interpretation. Another theory will require a different interpretation of ' causality.'

5

The causal anomalies in quantum mechanics show that the principle of contact-action must be avoided. The theory of elementary particles provides another argument. An elementary particle—

---

[1] F. Exner, *Vorlesungen über die physikalischen Grundlagen der Naturwissenschaft*, Vienna, 1922

in contrast to the classical particle—is characterised by definite values of mass and size and, moreover, possesses charge and spin. Obviously we cannot take the electron, say, as an elementary particle unless it is absolutely rigid. For a deformable body can always be regarded as being composed of smaller parts which slide against one another, and so its character of elementary particle is lost.[1] This argument suggests that within the dimensions of an electron action must be propagated with infinite speed. Or, the principle of action by contact which would ensure a finite speed in accord with relativity seems to be inapplicable in this sub-microscopic domain.

Recent discussion of radiation theory supports this argument. The problem is well known. When radiation is considered as due to an accelerated electron, the Coulomb term in the mathematical expression for the force leads to an infinite self-energy for the classical point electron. To obtain a correct value for radiation damping (so that this catastrophe is avoided) the advanced potential, or pre-acceleration, has to be introduced. This so-called absorber theory of radiation requires the use of advanced potentials only within the length of an electron radius, or within corresponding time intervals. 'All advanced effects . . . are limited to a time of the order of magnitude of the quantity (e^2/mc^3).'[2]

The use of advanced potentials violates the principle of action by contact, and so the classical model of causal action breaks down in the sub-microscopic domain. Of course this does not mean giving up the causal description of nature, only the attempt of representing causality by contact-action. The model of causality which requires a mechanism for transmitting the action is rejected and the idea of causal chain cannot be used in the description of sub-atomic events. Quantum mechanics has revised the basic concepts very much in this direction: we cannot speak of the path of a particle, and the classical concepts of position and momentum are applicable only with reservation.

6

If this is an acceptable interpretation of the present situation in physics, we have to renounce the use of the principle of action by contact. This does not mean that an action at a distance must be introduced in the theory of elementary particles; rather all description in terms of ' action ' appears to be useless. Field theory is then not applicable in this region, and no direct and simple fusion of field and particle concepts can be expected. But any new theory describing elementary particles must obey the requirement of ' correspondence ' and so pass into field theory, for the macroscopic region. No solution is known at present.

---

[1] A. March, *Natur und Erkenntnis*, Vienna, 1948

[2] Wheeler and Feynman, in *Reviews of Mod. Phys.*, 1945, 17, 157

It is, however, of interest to discuss the consequences entailed by a rejection of contact-action. Einstein has argued that the principle of action by contact is characteristic ' for the relative independence of spatially distant things. . . . Dropping this principle would make impossible the idea of closed system and thereby the formulation of empirically verifiable laws. . . . Without assumption of . . . mutual independence of existence of spatially distant objects (which originates from everyday life), physical thought is impossible. We cannot imagine how, without such clear separation, physical laws can be formulated or tested.'

This is an interesting argument and it illustrates the interrelations between the concepts of the physical thing-language. To give up the principle of contact-action completely, i.e. even in the macroscopic domain, would be too radical and there is no evidence to require it. A physical thing-language has been most useful in classical physics. This kind of language is, however, difficult to use in quantum mechanics, and new semantic rules are introduced to avoid the causal anomalies. Quantum mechanics, at least as we know it to-day, is a hybrid theory. It combines, in some instances, the account of microscopic and macroscopic phenomena, since quantum effects over large distances and during long time-intervals are described, provided action is small compared to h. But the description of sub-microscopic processes, e.g. of events within the size of an elementary particle, seems incompatible with the use of a thing-language in which particles represent the ' ultimate ' individuals. And the principle of contact-action *between* things cannot be used *at the same time* to describe events *inside* these things if they are to remain individuals. It appears to be preferable to drop the idea of action altogether, and this suggests a revision of the concept of velocity (already begun in quantum mechanics) and of the concepts of space and time. Indeed, many physicists feel that a good argument may be made in favour of a fundamental length, in spite of the failure of previous theories attempting to quantise directly space and time. The idea of closed system then does not apply to phenomena within this length which would be the limit of resolving events into space and time components, just as *h* is such a limit for conjugate variables in quantum mechanics. And is it reasonable to take a disintegrating atom as a closed system when, to do so, we must introduce unobservable particles such as the neutrino ?

The point I wish to make is this. A physical theory represents a semantic language-system, which, however, is not explicitly formalised and not even completely formulated. Instead a model is used. The *psychological* function of models in physical theory has long been recognised ; but their occasional misuse due to unwarranted ontological assumptions has brought them into disfavour. But a model is a rudimentary semantic system, that is, it provides an interpretation of the mathematical equations and so implicitly gives semantic rules for the use of basic concepts. This *logical* function of models is indispensable if no explicit semantic rules are given. The model underlying classical field theory makes the concept of causality refer to a certain process, e.g. contact-action and the inverse-square law. If this analysis is correct, then the principle of action by contact is merely a part of classical interpretation. The familiar description by a thing-language would still hold for larger distances but need not apply in the theory of elementary particles.

ERNEST H. HUTTEN

THE CONCEPT OF MIND AND THE UNCONSCIOUS

WHEN a flying stone breaks a pane of glass the hitting stone is usually called the cause, and the smash the effect. Since a steel plate would not break, the brittleness of the glass is part of the cause. From the relevant causal complex, hard flying stone, impact, brittle glass, and smash, the hit by the stone is usually, and appropriately, chosen out as the cause, and the breakage as the effect. Correlated causes and effects are usually thus selected from relevant causal complexes.

The striker cues a white billiard ball deftly down the centre of the table to return to the original spot. Though the presumed skill is exceptional, the example is legitimate. The cue-stroke is normally called the cause, and the precise return the effect. The cause is obviously selected, quite appropriately, from an elaborate causal complex. The table must be level, and its surface true. Constant practice has disciplined the nice physiological adjustment of the striker's arm. Item after item must be added to the relevant causal complex, for this also contains the nature of the stroke, the quality of the cue, the nature and perfect sphericity of the ball. It is not niggling to add the supporting floor, since the stroke could not be made, or the effect obtained, without it. It would be unusual to add the oxygen of the air, or the air itself, but it is not fantastic, in a full analysis, to add this essential condition of the whole event. Further analysis, though possible, leading, for instance, to final dependence on the sun, is not needed to point an important moral. However appropriately for usual purposes causes and effects may be chosen from relevant causal complexes, to ignore the essentially complex nature of causation may promote error. It seems to do so in *The Concept of Mind* [1] by Professor Gilbert Ryle.

Man has inveterately doubled himself into a body and an occupant soul. In one primitive notion a mannikin soul makes the body active. The 'double-life theory,' Ryle complains, is still rampant in philosophy. In effect, though Ryle does not so state it, the *mind*, as often conceived, is a highly refined version of the animistic mannikin occupant. As 'a ghost mysteriously ensconsed in a machine,' the body, it is a reminder of the mannikin. As ‘ a field of causes and effects,’ a mental analogue of the corresponding mechanical field in ‘ the human body,’ or in ‘ any other parcel of matter,’ it refines the naïvely conceived mannikin into a still animistically conceived entity. Ryle attacks ‘ the dogma of the Ghost in the Machine ’ by whittling away the complicated causal complex in the human person.

---

[1] London, 1949

A man meets a stranger and boasts to him. According to Ryle he has a boastful disposition, and the meeting makes him boast. The boastful inclination, or disposition, is a motive, but, not being an event, ‘ cannot be a cause ’ ; the meeting can be a cause, since it is an event. This seems to involve a curious consequence. A modest man does not boast when he meets the stranger. The same cause, the meeting, has different effects if it is the only causal item. The inclusion of the boastful disposition in the relevant causal complex removes the anomaly.

Ryle tries to cover the anomalous exclusion of causal efficacy from dispositions. He compares the disposition to the brittleness of glass : ‘ “ brittle ” is a dispositional adjective.’ Then he appeals to the ‘ law-like proposition.’ The glass broke because the stone hit it ; this explanation, according to Ryle, differs in type from explaining that the glass broke when the stone hit it because it is brittle. Though both statements contain a ‘ because,’ the first affirms a causal event, and the second states ‘ a law-like proposition,’ with an emphasis on the dispositional ‘ brittle.’ The latter involves a ‘ hypothetical proposition about the glass ’ which says, roughly, that ‘ the glass, *if* sharply struck or twisted, etc., *would* not dissolve or stretch or evaporate but fly into fragments.’ The brittleness of the glass, Ryle notes, is often called the ‘ reason ’ for the breakage, and analogously, the boastful disposition might be called the *reason* for the boastful outbreak. It is often convenient or appropriate to call a particular part of the effective causal complex the *reason*, especially a part which is neither an event nor mechanical, but such a reason is still causal. In the mechanical domain the victim tied to the stake to be burned would regard the binding rope as very causal, however he was assured that it was only the reason why he could not fly from the flames, or only exemplified the law-like proposition that if a man is firmly fastened he cannot move. The restriction even in mechanical causation to *events* involves causal anomalies. The level billiard table is as necessary for the skilled stroke as a deft cue or the absence of a sudden earthquake.

In *Paradise Lost*, during a battle with Michael's angels, the seraphs of Satan's host, applying flaming reeds to gunpowder in cannon, sent out a ' hail of iron globes.' The explosions in the cannon are what Sir Francis Bacon calls *magical instances*. In each instance the ' efficient agent,' the flaming reed, is meagre in comparison with the greatness of the effect, the explosive projection of the projectile. The causal efficacy does not reside in the flaming reed alone, as Bacon realises. It resides eminently in the gunpowder and its bestowal in the cannon. In Ryle's terminology, as brittle glass is dispositionally inclined to break when struck by a stone, or as sugar to dissolve when thrown into water, so the confined gunpowder has a disposition to explode when touched by the flame. The whole operative causal complex is ludicrously identified with the single event of this touch. Between the selected cause, the flaming reed, and the obvious effect, the hurtling explosion, lies the causal complex covered by the disposition of the gunpowder to explode. Ryle appeals to dispositions to shovel out the causes, but a boastful disposition is as much a causal complex, resulting in boastful outbursts under precipitating events, as confined gunpowder is a causal complex resulting in explosion at the touch of a flame.

Ryle is intent on dispelling ' occult ' causes, ' mental thrusts,' and ' ghostly causes of actions ' to depose the ghost from the machine. ' Mental processes,' he admits, do occur, though differently from ' physical processes.' The heroine in a novel kept one of her morning letters to read in private. We do ' know quite well why,' as Ryle says, for she had recognised her lover's handwriting. This reason is enough in ordinary life, but not for adequate analysis, since the recognition involves a large causal complex. Much had to happen to the heroine after her birth, for instance, before she could understand what writing is, or be able to recognise the hand. No doubt the farmer did come back from market with his pigs unsold because prices were too low, but the specified cause is picked out, appropriately enough, from an enormous causal complex. Much organised experience, for example, lies behind a knowledge of what prices are, or of how to calculate them. As a causal complex, elaborated out of organised experience, determines the deft billiard stroke, so it determines, or helps to determine, the heroine's secluded reading, the farmer's bootless return, or any event of ordinary adult life.

The simple understanding of any sentence in *The Concept of Mind* can, by convenient synecdoche, represent the elaborated causal complex that lies behind matured human understanding or action. Its meaning, in the last analysis, is mnemically organised experience, and a part of the whole unconscious that lies behind, and controls, conscious experience. Memory, essentially elaborately organised experience, constantly directs conscious thought and action, for the most part unconsciously, though it does result, from time to time, in consciously explicit recognition. Any simple perceptual experience, such as recognising an object to be a tree, is determined by mnemically organised experience, largely spontaneously automatic, and predominantly unconscious. When the batsman raises his bat to strike, the overt action depends, apart from his intention, on a physiological complex. Any ostensibly mental process, such as understanding the argument in *The Concept of Mind*, depends upon an elaborate organisation of experience in the unconscious. The causes underlying any action, such as kicking a football, or any admittedly mental process or performance, such as reading *The Concept of Mind*, are mainly ' occult ' : they are not observed, and when recognised at all, are successfully probed or reflectively inferred, though the whole relevant causal complex is only exposed with differing degrees of completeness. The mnemic organisation of experience in the unconscious mind is ultimately enigmatical, however justly it is inferred or its structure partly divined.

To follow a familiar tune, knowing how it goes, Ryle writes, ' a set of auditory expectation propensities ' must have been acquired. ' Memories of past hearings of the tune need not well up, or be called up.' These past hearings, however, have their cumulative consequence in a mnemically organised experience. The tune will not be familiar to one who has not heard the tune before, and will be so to one who has. The hearing obviously goes on differently in the two people, as Ryle notes, but ' specially schooled hearing ' is a very lean explanation. The schooling lodges some causal efficacy to produce the familiarity. There is not ' something extra going on ' in the second, and schooled, person, Ryle affirms, or ' a report of extra occurrences.' There is something in the mnemically organised unconscious, probably a process during the hearing, to *cause* the familiarity. The ' auditory expectation propensities ' are depleted of causality, as all other dispositions are depleted in *The Concept of Mind*. Ryle has good reason, in defending his thesis, for stripping away the mnemic organisation of experience in the unconscious mind. This complex causal field is more like an animistic occupant than the ' Ghost in the Machine ' should be if it is to be easily expelled. Very characteristically, *The Concept of Mind* virtually ignores the unconscious mind, that elaborate causal complex mnemically organised out of experience. The mannikin, indeed, seems to be still dimly discernible in the highly reconstituted version presented by the modern concept of the unconscious. The mind, on this modern interpretation, is at least more animistic than Ryle relishes.

Ryle's thimble example illustrates how he ignores the unconscious. The finder of this useful article must recognise it ; to recognise it he must have learned what thimbles ' look like,' and he must not have forgotten what he has learned. This, stated by Ryle, involves a greater complexity of process than he admits. The proposition ' he espied the thimble,' he affirms, ' has a considerable logical complexity,' but the ' considerable complication of processes ' involved in the espying is a myth. This, curiously enough, contrasts the complexity of the logical process with the simplicity of the mental process that underlies all the logical. If ' the thimble is obvious to the observer ' when he has his visual sensation of it, there need be ' no extra thinking or pondering,' and ' no puzzlings or reminiscences need be performed.' No ' swift and unnoticed inferences,' or ' sudden and unrememberable intellectual leaps ' need occur unwitnessed in the ' finder's stream of consciousness.' It is the office of the mnemically organised unconscious to promote the rapid spontaneous recognition of the thimble, without obtruding itself, and dispensing with ratiocination. Such perceptual judgments are intuitive : they are unanalysed and direct. If the content is analysed, the tale of relevant experiences on which the quick recognition ultimately depends is complex.

The ' having a visual sensation in a thimble-seeing frame of mind ' ignores the operative unconscious which makes the ' frame of mind ' intelligible. The observer does not puzzle out the significance of the visual appearance by conscious reference to similar past appearances, and by connecting them, for instance, reminiscently with fingers and needles. There is inference from visual appearance to functional thimble, but not in the stream of *consciousness*. What would otherwise be conscious inference, involving conscious ' reminiscences ' and, possibly, conscious ' intellectual leaps,' has a functional equivalent in unconscious process. The observer sees the visual appearance, and past experience, mnemically organised, operates unconsciously to promote the conscious recognition. The unconscious process relieves consciousness of as much ratiocinative trouble as it can. The observer is, normally, not aware that he is recognising the thimble, as he might be if he had to think out his recognition. He simply recognises, or is ' thimbly ' recognitive. The observer can effectively recognise the thimble without being aware that he is recognising it, and his unconsciously prompted recognition omits the needless extra awareness. Here, as elsewhere, Ryle ignores the unconscious and, in so doing, cuts from the mind ' a monstrous cantle out.' This reduces the ghost, but makes his residue, if there is a residue, unintelligible.

JOSHUA C. GREGORY

DISCUSSIONS

The Law of Inertia

In his paper 'On the Foundations of Dynamics' (this *Journal*, 1950, 1, 92), Dr Whitrow discusses whether the Law of Inertia 'should be regarded' as self-evident, empirical, *a priori*, etc. Dr Whitrow's preferred interpretation of the Law as 'defining jointly with the other laws of motion, a method of analysing motions' may express the most valuable way of using it in the present state of knowledge.

But there is no one way in which the Law 'should be regarded.' The logical or other status of a principle depends on the manner in which it is used and can only be discussed *in relation to a specified context*. Thus the Law of Inertia can appear with equal legitimacy in many forms : e.g. as an axiom, a theorem, or as a methodological postulate forming part of a method, using assumptions not yet fully clarified, for analysing observed motions.

L. L. Whyte

Mr L. L. Whyte argues that the Law of Inertia has a meaning only in a specified context. My point was to discover how the law can best be regarded in practice by those who use it for the investigation of the physical world. It has been suggested in the past that this law is an *a priori* truth, an empirical proposition, a convention, etc. I examined these suggestions in turn and argued that the most satisfactory solution was to regard it purely as a tool for, or method of, analysing motions. It is invented by man, but has an empirical aspect because it is found to be a means whereby we can analyse motions in the external world. Any law of motion is of this character and has its own field of application. Newton's laws of motion can be usefully applied to a wide, but not universal, range of natural phenomena.

G. J. Whitrow

Note on Philipp Frank's Interpretation of Science

In his paper on 'Metaphysical Interpretations of Science' (this *Journal*, 1950, 1, 60) Dr Philipp Frank gives an interesting interpretation of the history of science and philosophy, the latter being regarded as composed of metaphysical interpretations of science : ' " Science " and " Philosophy " . . . or, to use the language of this paper : science and its metaphysical interpretation . . . ' (ibid. p. 67). This interpretation must partly be ascribed to Frank's conviction that metaphysics is meaningless : ' If we apply to metaphysics the criteria of truth which have been generally accepted in modern science, we can conclude, on good grounds, that tenets of metaphysics . . . are neither true nor false but meaningless ' (ibid. p. 60). It is the purpose of this Note to draw attention once again to the fact that the criterion of truth by which metaphysics is reduced to meaninglessness is itself metaphysical, in which case Dr Frank's interpretation must be treated with some reserve.

Frank's notion of what is metaphysical may be obtained from the following quotations : ' If no statement about observable facts can be derived we say that these principles belong to metaphysics ' (ibid. p. 61), and ' If a statement which claims to give a " metaphysical implication of science " can be checked by the agreement of its consequences with experience we have to regard these would-be metaphysical implications as part of science ' (ibid. p. 62), and again ' we shall understand . . . by a " metaphysical interpretation of science " a search for the reality behind the physical phenomena ' (ibid. p. 62). It is perhaps because of Frank's preoccupation with this last definition that his attention has been diverted from the metaphysical nature of the verification principle of logical empiricism.

A further quotation will be of some use : ' Planck started from the obvious fact that every physical theory contains two types of terms : the terms denoting sense observations (red, warm . . .) and the symbols forming the general principles (potential, force . . .). The . . . ' positivist ' . . . claims that beyond these two realms of concepts nothing can be said about physical science ' (ibid. p. 62). It is clear from the text that Frank agrees with this formulation. Granting to physical theory these two types of terms and these two only, what status can be given to the verification principle of logical empiricism ?

Frank's verification principle, implicitly contained in the above quotations, may be put thus : A statement has meaning only if its consequences may be checked by agreement with experience. What status then has this principle ? If meaningful, its consequences must agree with experience but no one has suggested how this might be tested, so even if it could be regarded as a scientific hypothesis it must be adjudged ' not proven.' But how could it be regarded as a scientific hypothesis when it does not contain the two types of terms as described in the last quotation given above? It is in fact of no use to anybody unless it is considered in conjunction with a particular theory of science. Hence it cannot be regarded as empirical. Since neither Frank nor other logical empiricists suggest that the principle is tautological, we must conclude that it is metaphysical. Of course, A. J. Ayer wishes the principle to be regarded as a definition (*Language, Truth and Logic*, 2nd edn., 1946, p. 16). ' Thus, while I wish the principle to be regarded, not as an empirical hypothesis, but as a definition, it is not supposed to be entirely arbitrary.' But if it is not arbitrary, it must be considered as having empirical reference and therefore if we are right in suggesting that it cannot be empirical it would seem necessary to regard it as metaphysical.

Now, if metaphysics cannot be reduced to meaninglessness on the ' good grounds ' put forward by Dr Frank, it is questionable whether metaphysics may universally be interpreted in the way suggested by him.

OWEN POTTER

REVIEWS

THEORIES OF RELATIVITY

THE publication in the *Library of Living Philosophers* of a monumental volume devoted to the work of Einstein [1] provides an excellent opportunity for a reassessment of the significance of his outstanding contributions to theoretical physics and natural philosophy. Running through most of the twenty-seven essays in this book, which begins and ends with contributions by Einstein himself, are two master-themes : Einstein's creation of modern relativity theory, and his ultimate rejection of modern quantum theory. Both of these tend to place him in lonely eminence. Whereas the construction of quantum theory is a supreme illustration of the co-operative character of science, the creation of relativity theory is an equally impressive instance of the contrary thesis, namely that outstandingly original advances are usually the work of isolated individuals. Again, whereas most contemporary theoretical physicists accept the indeterministic philosophy associated with quantum mechanics, Einstein rejects this philosophy from the very depths of his being. This rejection goes much deeper than Newton's rejection of the wave-theory of light, to which his contemporary Huyghens made so great a contribution. Newton rejected the wave-theory because it seemed to him incapable of explaining all the known facts. Einstein rejects quantum mechanics because it appears to him to be diametrically opposed to his conception of explanation itself. His difficulty is epistemological and springs from his philosophy of science.

In his short contribution to another volume in the same series—one devoted to the work of Bertrand Russell—Einstein maintains that ' The present difficulties of his science force the physicist to come to grips with philosophical problems to a greater degree than was the case in earlier generations.' The more fashionable empiricist attitude, however, is a reaction against the ' metaphysical ' beliefs of the past, a reaction which is all the more impressive and compelling because of the overwhelming increase in our understanding and control of natural phenomena which followed the breakdown of those beliefs. In a famous passage, John Locke spoke of himself as an under-labourer ' clearing the ground a little and removing some of the rubbish that lies in the way to knowledge.' He believed that a proper understanding of nature is possible only when we empty our minds of all preconceived ideas. Nevertheless, although it may not be true that nature abhors a vacuum, our minds cannot begin to function from mere ' emptiness.' As Kant wrote at the beginning of the *Critique of Pure Reason*, ' though all our knowledge begins with experience, it does not follow that it all arises out of experience.' In similar vein, Polanyi, in a recent lecture on *The Nature of Scientific Convictions*, argued incisively against the positivist belief in the existence of a core of hard facts or incontestable primary sensations which any theory must accept as such. He showed how difficult it is to discover any primary sensations which are given previous to our interpretation of them, and concluded that there is always a measure of choice in our manner of perception.

---

[1] *Albert Einstein : Philosopher-Scientist*, edited by Paul Arthur Schilpp (*The Library of Living Philosophers*, vol. vii.) ; Evanston, Illinois, The Library of Living Philosophers, Inc. 1949 (London, Cambridge University Press, 1950). Pp. xvi + 781. 75s.

Although not subscribing to all Kant's views on the nature of space, time, etc., Einstein has adopted an essentially Kantian attitude in his belief that ' the concepts which arise in our thought and in our linguistic expressions are all—when viewed logically—the free creations of thought which cannot inductively be gained from sense-experiences.' In particular, Einstein would certainly agree that scientific research is itself an art, and that like all the other arts its life-blood is the human power of imagination. Although this conviction is one which many men of science share—indeed W. I. B. Beveridge, the author of a recent book on *The Art of Scientific Investigation*, points out that imagination rather than the study of formal treatises on scientific method is the indispensable factor in the practice of this art—nevertheless, few men of science today would agree whole-heartedly with the general point of view expressed by Einstein in the sentence quoted.

In his Herbert Spencer lecture delivered at Oxford in 1933, Einstein drew particular attention to the fictitious element in the basic concepts of physics. Newton, the architect of the first workable system of theoretical physics, believed that its foundations were entirely derived from experience. However, ' one can see,' said Einstein, ' from the way he formulated his views that Newton felt by no means comfortable about the concept of absolute space, which embodied that of absolute rest ; for he was alive to the fact that nothing in experience seemed to correspond to this latter concept. He also felt uneasy about the introduction of action at a distance.' Nevertheless, the practical success of Newton's theory was such that during the eighteenth and nineteenth centuries it was generally believed that the basic concepts and laws of physics were derivable by a logical process of abstraction from experiments. Einstein then went on to argue that the incorrectness of this view was first clearly demonstrated by the general theory of relativity. ' For this theory revealed that it was possible for us, using basic principles very far removed from those of Newton, to do justice to the entire range of the data of experience in a manner even more complete and satisfactory than was possible with Newton's principles.' Thus the fictitious character of the principles is made obvious by the fact that two essentially different theories both lead to a large measure of agreement with experience. One recalls Pasteur's remark that the hallmark of a good scientific theory is not its ' truth ' but its ' fruitfulness.'

In this analysis of the nature of physical theories, I believe that Einstein's view is essentially correct, but does it necessarily commit us to his belief that the fundamental concepts of physics are ' freely chosen ' by us? Although Einstein has been influenced by Kant, this particular belief appears to stand in marked contrast both to Kant's own conception of the unique *a priori* character of space, time, etc., and to the crude empiricist view that the true laws of nature are uniquely given by experience ; in both these philosophies a unique solution is imposed on us. On the other hand, Poincaré stressed the conventional character of our basic concepts and hence the plurality of theories. Einstein, however, believes that by its free activity the mind can construct a particular formal theory which is better adapted to experience than any other and so is, in this sense, a true and unique representation of reality. It is well known that he believes that the basic concept of such a theory is the field concept introduced into physics by Faraday and Maxwell, a concept which avoids the Newtonian dilemma concerning action at a distance. During the last forty or so years he has striven to adapt this concept so that it gives a unique and universal description of physical reality. Closely linked with this concept are the ideas of ' continuity ' and strict ' causality.' As Rosenfeld has recently pointed out, Einstein attempts to exalt his conception of the field to the rank of a universal pattern for the description of nature, and discounts as a temporary imperfection the failure of this pattern to apply to atomic phenomena. Thus, believing simultaneously in the free nature of our concepts and in the existence of a unique picture of reality, Einstein is forced to admit that the harmony which he seeks is incomprehensible, an ultimate mystery of the co-existence of mind and matter.

Since the time of Galileo it has become clear to all students of natural philosophy that the most fruitful starting point is the study of motion, and classical physics, relativity physics and quantum physics are all essentially theories of motion. There are two ways in which we can regard the successive development of these theories, namely by concentrating either on the phenomena to which the theories are applied or on the foundations of the theories themselves. With regard to the former, Einstein remarks in his stimulating autobiographical introduction to the present book that however evident may be the demand that a successful theory must not contradict empirical facts ' its application turns out to be quite delicate. For it is often, perhaps even always, possible to adhere to a general theoretical foundation by securing the adaptation of the theory to the facts by means of artificial additional assumptions ' (p. 21). Einstein, therefore, attaches particular importance to the foundations of a satisfactory theory, arguing that ' we prize a theory more highly if, from the logical standpoint, it is not the result of an arbitrary choice among theories which, among themselves, are of equal value and analogously constructed.' He states in parenthesis that he is speaking only of theories ' whose object is the *totality* of all physical appearances ' (p. 23). A few pages later, however, he tells us that in his opinion classical thermodynamics is ' the only physical theory of universal content concerning which I am convinced that, within the framework of the applicability of its basic concepts, it will never be overthrown ' (p. 33), thereby apparently implying that some restriction should be placed on the interpretation of the word *totality* occurring in the previous passage. Nevertheless, the aim of Einstein's researches in physics has been the creation of a *universal* theory of motion applicable to all physical phenomena. In this quest of an all-embracing theory Einstein has long departed, as Philipp Frank points out (p. 274), from Mach's positivistic requirement that the laws of physics should be formulated so that they contain only concepts which could be defined by direct observations or at least by a short chain of thoughts connected with direct observations. This requirement was met by the special theory of relativity, the ' operational ' character of which is stressed by Bridgman in his essay on " Einstein's Theories and the Operational Point of View ". The special theory was Einstein's greatest achievement as a young man when it was not yet clear to him that ' the approach to a more profound knowledge of the basic principles of physics is tied up with the most intricate mathematical methods ' (p. 17). The creation of the general theory, however, inspired the faith of his mature years that ' we can discover by means of purely mathematical constructions the concepts and laws connecting them with each other, which furnish the key to the understanding of natural phenomena ' (p. 283). Consequently, although experience must still be the sole criterion of the physical utility of a mathematical construction, ' the creative principle resides in mathematics.'

Einstein thus takes his place high among the great mathematicians who have sought in some mathematical concept the master key to unlock the secrets of the physical universe. Lagrange sought this key in his idea of the energy function, and Hamilton in his invention of the characteristic function. The grandeur and power of these conceptions, however, cannot conceal from us that they function as tools of research adapted for definite types of investigation and not as universal ' open sesames ' to the whole of physical reality. Einstein's belief in the existence of a master key springs from the peculiar epistemological character which has come to be attached to his general theory of relativity. Even so acute a critic of Einstein's philosophy of science as Max Born—whose contribution to the present volume on " Einstein's Statistical Theories " concludes with his rejection of Einstein's faith in ' perfect laws in the world of things existing as real objects which I try to grasp in a wildly speculative way '—cannot withhold his admiration from Einstein's achievement in creating a field theory of gravitation which is 'logically perfect.' Nevertheless, this theory has its limitations.

First, it is well known that so far general relativity is not regarded as a mathematical technique which all students of physics must master. Apart from the three original observational tests, the theory as a whole has appealed to the pure mathematician rather than to the theoretical physicist. In this respect it stands in marked contrast to the special theory. Moreover, if general relativity is a first approximation to a universal theory, then the latter may prove to be of such mathematical complexity that few if any of us will be able to adapt it to special problems. Thus, what is gained in 'epistemological simplicity' may well be forfeited in the lack of practical applicability.

This criticism of the mathematical technique, however, should not be over-emphasised, particularly as the history of mathematics presents many examples of the ultimate simplification of unwieldy techniques, e.g. the Sumerians and the Babylonians would be amazed at the ease with which we now multiply numbers. A second and more fundamental limitation of the general theory of relativity concerns the very citadel of its strength, namely its 'epistemological simplicity.' It is instructive to recapitulate very briefly Einstein's object in creating this theory.

In the decade immediately following the publication of the special theory in 1905, Einstein pursued the dual object of extending the relativity principle to all relative motions and the field concept to all types of force, in particular to gravitational forces. His genius was perhaps never more evident than in his realisation that the identity of inertial and gravitational mass—which had been checked experimentally to a high order of accuracy by Eötvös—was not something to be taken for granted, as others had unconsciously presupposed, but demanded elucidation. The special theory had shown that inertial mass depends on relative motion. Accepting this conclusion, it was then necessary to consider the effect of motion on the classical law of gravitation. It appeared to Einstein that the problem was insoluble within the framework of the special theory. 'Now it came to me,' he writes (p. 65), 'The fact of the equality of inert and heavy mass, i.e. the fact of the independence of the gravitational acceleration of the nature of the falling substance, may be expressed as follows: In a gravitational field (of small spatial extension) things behave as they do in a space free of gravitation, if one introduces in it, in place of an "inertial system," a reference system which is accelerated relative to an inertial system.'

To illustrate this point, Einstein devised a celebrated argument concerning a man in a lift. He imagined a lift at the top of an indefinitely high skyscraper. The cable breaks and the lift falls freely. The man in the falling lift drops objects inside the lift and finds that he can ignore the gravitational field because relative to him objects are unaccelerated when he drops them. Thus all bodies in the lift behave in the way expected from the law of inertia. To an observer outside the lift who can look into it, the objects therein appear to move in accordance with the law of gravitation, but a physicist whose whole life is spent in the falling lift would believe himself to be in possession of an inertial system and would refer all laws of motion to this system. It would be natural for him to regard the lift to be at rest and the seat of an inertial system. As is well known, Einstein went on to argue that there is no ultimate reason why a frame of reference which is accelerated with respect to a given inertial frame should be less fundamental than one which is unaccelerated. Consequently, all frames of reference in all possible motions must be on the same footing and the laws of nature are to be expressed by equations which hold good for all frames.

The type of equation for which Einstein looked was a field equation, and he sought the simplest equations relating the structure of space-time to the distribution of matter and energy by means of a non-uniform Riemannian geometry. His method involved certain arbitrary features which were justified by their success in accounting for the perihelion motion of Mercury, etc., but which it was soon realised might require some modification if the theory was to be extended to include electromagnetic forces. The general principle of covariance, however, seemed to be immune from any such tampering as it was the solid rock on which the whole theory was securely based.

The epistemological cogency of the principle of covariance was due to the fact that it appeared to be the most 'natural,' i.e. least arbitrary, way of expressing the idea that all motion must be relative. Belief in this idea as the foundation stone of dynamics has been instilled into generations of students. To cast doubt on it seems to be equivalent to questioning the whole framework of theoretical physics which has been built up since Galileo and others overthrew the ancient belief in absolute motion. Nevertheless, in common with so many ideas which have played a decisive role in history, it may be a 'myth.'

The ancient philosophers, for example Plato and Aristotle, regarded motion *cosmologically*. They always had the general background of nature in mind. They have been criticised for this, but we are now beginning to realise that, although they 'looked through a glass darkly,' they were in some respects much nearer the heart of things than used to be thought. Fundamental problems of motion cannot be completely divorced from our general philosophical outlook. For example, if the Michelson-Morley experiment had been performed in 1587, instead of 1887, it would probably have been interpreted as conclusive evidence of the immobility of the Earth. Instead, it actually became the basis for a great extension of the idea of the relativity of motion. Similarly, criticism of the belief that all motion is relative arises from the new ideas which we are now forming concerning the ultimate background of the physical universe.

Consider once more Einstein's argument of the man in the freely falling lift who imagines that he is in a simple inertial system, whereas, to a man outside, the lift is in a gravitational field. Einstein assumes that the man outside the lift can see what is going on in the lift, but the man inside the lift evidently is not allowed to look out—or at least, it is tacitly supposed that this consideration would be irrelevant. Now such a consideration is certainly irrelevant so long as we regard space as a whole to be effectively empty. In applying the technique of general relativity to the problem of Mercury, for example, we presuppose that the system concerned is silhouetted against a dark empty background and against such a background it is ' natural,' or ' epistemologically primitive ' to impose the condition that all frames of reference are equivalent. But if, following recent trends in observational knowledge, we cease to regard the background as a vacuum, then the background itself may tend to isolate a privileged class of frames of reference, namely all those at local rest relative to this background. Motion with respect to such frames would then be the counterpart of the absolute motion of the ancients.

The question of whether motion is ' relative ' or ' absolute ' does not refer to nature but rather to alternative concepts of nature. Are we to regard these alternatives as free concepts ? With his eye directed to quantum physics, Rosenfeld has argued that although Einstein is fully justified in regarding the basic concepts of physics as creations of the mind he is mistaken in claiming that they are *free* concepts. Similarly, with an eye directed to cosmological physics, we may argue as follows. If the concepts of an empty and of a non-empty background were free concepts then there would be no reason why we should not always choose the former and with it the general relativity principle. I believe, however, that in choosing one concept rather than another we are like a painter who chooses, say, water-colours rather than oils, thereby restricting the effects which he can achieve. Thus, it is a fact of experience that choosing the technique of general relativity Einstein can account for the perihelion motion of Mercury, presumably because this problem readily lends itself to the appropriate type of mathematical idealisation ; but it does not automatically follow that the same technique (appropriately adapted) will be as uniquely successful in analysing the motions of the nebulae.

To sum up : theories of relativity like all theories of motion are methods which are devised for the study of physical phenomena. We do not value a theory if it is merely an *ad hoc* technique for dealing with a single situation, but only if it has a wide range of applications. Moreover, as Einstein rightly points out, we cannot remain content with a theory which does not impress us as epistemologically ' satisfactory.' Our physical concepts, however, are not completely free, for although it is true that 'there is no way from experience to the setting up of a theory' (p. 89), experience remains the ultimate criterion which restrains the imagination. Even if the search for a unified theory ultimately yields rich treasure, it is inspired by the illusion that we can form in our minds a unique mirror image of physical reality. The change now taking place in our ideas concerning the background of nature challenges the principle which was thought to be the necessary foundation of such an image.

G. J. Whitrow

Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Ritual, Susanne K. Langer, Harvard University Press and Oxford University Press, 1942, pp. xiv + 313 (out of print); Mentor Books, New American Library, N.Y., 1948, pp. 248. $0.35 (in print; unabridged cheap edition, with index, notes, but short reading list in place of original bibliography).

Readers of this Journal may be glad to have their attention called to this nine-year-old book, and to the following instructive facts.

Philosophy in a New Key was issued by Harvard in 1942, reviewed favourably in the *Times Literary Supplement* (31-7-43), and highly praised by Professor L. A. Reid in an essay-review in *Mind* (Jan. 1945) as of 'exceptional interest and merit.' The Harvard edition went out of print, few copies having reached this country. The Mentor edition (offspring of Penguin, now independent) appeared in 1948, and over 80,000 copies have been sold. But owing to war conditions, lack of interest by publishers, and the fact that the book does not fall within any one of the dominant 'subjects,' such as symbolic logic, philology, philosophy of art, religion, etc., there has as yet been no United Kingdom edition. This has had the unfortunate result that, though known to a handful of specialists, this scholarly, original and delightful work has not yet reached the wider circle of philosophical, scientific and general readers outside the United States. Those who are left unsatisfied by the present cursory description can turn to the *Mind* essay for a fuller analysis of the argument, or order the book for the price of twenty cigarettes.

Mrs Langer is a pupil of A. N. Whitehead, who has passed from the study of symbolic logic to the wider problem of the role of symbolic forms of all kinds in the origins of human culture. Recognising that the study of symbols in language and mathematics may be too narrow to yield full understanding of their nature, she has sought to develop a general philosophy of symbolic forms in relation to the development of human faculties. Her aim is to realise Cassirer's conception of a grammar of the symbolic function covering all its special expressions and idioms as displayed in Language and Art, Myth and Religion. This aim goes beyond but does not conflict with symbolic logic and linguistic studies, provided that those special fields do not claim to cover the entire realm of symbolism, semantics, meaning and knowledge.

With a friendly salute to the great achievements of the school of Russell, Wittgenstein, Ogden and Richards, Carnap, etc., and all the students of language, which has tended to restrict ' symbolism ' to the study of the influence of linguistic symbols on thought and has started off from the concept of logical form expressed in such symbols, Mrs Langer has gone back to a prior and more general issue : the symbolising tendency or function evident in all culture, including the special forms which are separated (at some risk to understanding) in the categories of Art, Music, Myth, Religion, Language, Philosophy, Mathematics, and Science. The ultimate objective is a general theory of all characteristically human activities as the expression of a universal symbolising tendency in *Homo sapiens* (though one must not forget the use of non-linguistic symbols by bees and other animals), and here her main predecessors and teachers are, as she states, Whitehead and Cassirer.[1] But she has known how to stand on their shoulders and to see much further than they. A second reading of the book has left one reader in no doubt that he is watching a constructive advance in thought. Mrs Langer appears to understand better than any other writer known to him the essential character of the symbolic transformation of experience (her term), more precisely the universal human and psychological factors underlying the origins and growth of all cultural activities, as against the special forms characteristic of particular cultures. This is a personal view ; since *Philosophy in a New Key* is an essay in speculative philosophical anthropology the final judgment on its scientific value may come from the anthropologists, the social scientists, or the biologists of man.[2]

The essence of Mrs Langer's theme is that the term ' Symbolism ' is appropriate to describe a basic need in man which is also the essential act of mind: the symbolic transformation of experience.[1] Symbolism is expression (in Langer's new sense) and this is an end in itself (e.g. the child's or primitive's delight in naming, talking, etc., even to himself), and not merely an instrument for communication or survival, as orthodox epistemology has wrongly assumed. The elimination of metaphysical pseudo-questions by modern logic is valid, but not the contemporary logicians' unduly narrow concept of ' meaning.' Language is *not* the only means of articulating thought, and everything which is not speakable thought is *not* formless feeling. For gesture, ritual, art, music, and poetry all express (and not merely convey) articulated meaning, that is the morphology of aspects of mental life: patterns of motion and rest; tension and release; preparation, fulfilment, excitation; sudden change; in fact all the stationary forms and dynamic patterns of human experience. Music most clearly symbolises this morphology of feeling, and invites not emotional response, but insight; music expresses the composer's knowledge of this morphology, and incidentally may share his knowledge with others, but music does not evoke primary emotions when it appeals to the aesthetic sense. Thus discursive symbolism is not the only vehicle for meaningful thought. The function of a symbol is to represent or connote the structure of a series of similar experiences; the symbol fixes what is common to them and makes it available for recall, thus enlarging experience.[2] Feelings have definite forms which in the course of human development become progressively articulated, and aesthetic, religious, and other cultural symbols represent and express this articulation. All forms of symbolisation spring from the basic human functional tendency to symbolise, which arises from the properties of neural organisation in the brain. This tendency cannot be derived from directly adaptive factors; communication is only one of its many consequences, which often run counter to biological value. The new key in philosophy is the movement of the last thirty years (better, fifty) towards an understanding of the primary role of symbolism in human affairs.

---

[1] See also L. A. Reid, *Knowledge and Truth*, 1923, Chapter IX on Non-Propositional Truth.

[2] See Charles Morris, *Signs, Language, and Behaviour*, 1946, p. 189, for names of other workers in the same direction. But Morris's definition of a ' sign ' is narrower than Langer's concept of ' symbol.'

Mrs Langer's central conception of the ' symbolic transformation of experience ' can only be appreciated by reading the book. She uses the term ' symbol ' in an enlarged sense which can include as special applications the usage both of the logicians and of Freud and Jung. This is a consequence of the fact that her approach is balanced and historical rather than either idealist-linguistic-logical or psychological-analytical. It may be that this deeper analysis of the nature of symbolism can serve to clarify and strengthen the foundations of logic by revealing how the mind acquires or creates the material with which logic starts and what relation logical mental processes bear to other kinds.

---

[1] Compare Schleiermacher (Sittenl. § 151), who suggested that there is present in all cognition a symbolising activity of the understanding, 1835!

[2] See G. Ferrero, *Les lois psychologiques du symbolisme*, 1895: ' The function of a symbol is to evoke certain states of awareness.'

No suggestion can be given here of the richness of Mrs Langer's argument. She is perhaps strongest on music and weakest on the visual arts. The writing is sober, clear and logical even when treating aspects of human experience which are symbolised in non-linguistic forms.

One caution may be desirable, supporting a similar comment by Professor Reid. Symbolisation is surely not an irreducible process. It must be the expression of properties of the brain, which automatically bring about the maturing of faculties latent in man's hereditary constitution by a progressive adjustment of neural organisation to the facts of the universe, including those of man's own make-up. Symbolisation is a consequence of properties of brain organisation expressed in a mode of transformation of brain activity patterns (not yet identified by electro-physiology), the effects of which transcend the mere adaptive maintenance of human life as it already is at any moment, because new potentialities of these brain properties are still in course of being evoked. The satisfaction in naming, thinking, talking, composing, inventing and creating is the organism's sense of fulfilment in discovering new ways of expressing its own vitality and in developing capacities previously unexploited.

'I would urge philosophers to read this most important book, which is far richer and more penetrating than any bald summary can indicate' (Reid, 1945), and not only professional philosophers, but all who seek a broader view of the origins and foundations of culture and of knowledge.

L. L. Whyte

Les Fondements Logiques des Mathématiques, E. W. Beth, Gauthier-Villars, Paris, 1950. Pp. 222. Frs. 1400.

This work begins a new series of monographs which is to be published under the title *Collection de Logique Mathématique*. The editor is to be congratulated on having secured, as author of the first volume, the distinguished Dutch scholar whose article on 'Critical Epochs in the Development of the Theory of Science' appeared last May in this *Journal*.

The traditional doctrine regarding science may be traced backwards to Aristotle, who taught that science can be established deductively by starting from certain irreducible postulates, which cannot themselves be proved, and are to be grasped intuitively ; they cannot be overthrown, or even modified or limited, by any result of scientific investigation. Among them it has been customary to include the laws of logic and (until the nineteenth century) the axioms of geometry.

The classical example of a science having this character was mathematics, the fundamental principles of which were borrowed by Aristotle chiefly from Eudoxus and Theaetetus and were followed systematically to their consequences not long afterwards by Euclid. Thus it came about that an inquiry as to the nature and origin of the intuitive truths was bound to examine in the first place the foundations of mathematics.

It was, however, not until the latter part of the nineteenth century that real progress was made in this investigation. In the 1880's Frege published his epoch-making work on the nature of number, and Cantor made his researches on aggregates and the continuum ; and in the following decade Peano of Turin devised an ideography—a representation of logical ideas by symbols—which was capable of replacing ordinary language altogether for the purposes of precise reasoning in the philosophy of mathematics. This tool began to be used in 1900 by Whitehead and Russell, who by its means found that the whole numbers 1, 2, 3, . . . can be defined in terms of purely logical concepts (such as *or*, *not*, *class*), and that all pure mathematics can be built up on this foundation ; so that *mathematics is actually part of formal logic.* It was admitted that for mathematical purposes certain axioms must be added to those that are usually found in treatises on logic, e.g. the axiom that the universe contains an infinity of individuals : but this extension of logic did not affect the main position.

Russell's discovery was not received with universal assent, partly on account of some difficulties which were brought into prominence by his work. These were the *antinomies*, eleven of which are discussed in the book under review. The oldest is that of Epimenides, who said, ' I am lying ;' it is easily seen that the alternative possibilities as to what he was really doing both lead to contradictions, the complete explanation of which was found by Tarski only in 1930. The antinomies led some mathematicians to feel that mathematics was discredited by Russell's identification of it with logic ; and Hilbert proposed an alternative philosophy of mathematics which he claimed to be entirely free from contradiction : this has, however, been generally abandoned since Gödel proved in 1931 that some propositions of mathematics cannot be either proved or disproved by any system based on axioms such as those of *Principia Mathematica.*

A wholly different school of opposition to the Russell doctrine was that of the *intuitionists* and *finitists.* They abandoned the attempt to justify large parts of traditional mathematics ; in particular they rejected all proofs by *reductio ad absurdum*, and all propositions involving infinite collections or infinite series. The disastrous consequences to mathematical analysis of adopting such a position has prevented it from gaining any general acceptance, but it is not easy to disprove.

Professor Beth's work gives an account of the main line of progress and of all contributory movements. The researches of Carnap and others during the last twenty years on syntax and semantics, and those of Church, Quine, and the writers in the *Journal of Symbolic Logic* generally, are described, with ample bibliographical notices. Altogether an excellent book.

EDMUND T. WHITTAKER

Analytical Biology, G. Sommerhof, Oxford University Press, 1950. Pp. viii + 208. 17s. 6d.

Mr SOMMERHOF's aim is very ambitious. He wishes to make a critical analysis of the nature of life in terms which will be ultimately expressible in the 'language' of physics. In fact he is mainly concerned with its apparently purposive or goal-directed character. The core of the book is a mathematical formulation on pp. 89–92 of the following problem. A class of events S occur at time t_0. These cause at time t_1 on the one hand a set $E^1, E^2, E^3, \ldots$ of events external to the organism, and on the other a set of responses $R^1, R^2, R^3, \ldots$ by it. These are so 'directively correlated' that a goal G is reached at time t_2. For example an animal sees an object which is moving at time t_0, responds to it by appropriate motions at time t_1, and catches it at time t_2. The author is able to state the conditions for success, in a highly abstract form, as a set of partial differential equations. If these are fulfilled, G is achieved ; and the wider the range of S the greater the degree of directive correlation.

Unfortunately, he makes the following statement (p. 96, his italics), '*The reader who believes that a statistical approach could assist in the solution of our present problem has misunderstood the fundamental nature of the problem.*' Either the reviewer or the author has done so. As I see it the problem is essentially statistical. A bird, under the influence of hormonal urges and the visual stimuli afforded by small objects such as twigs and leaves, builds a nest G. The nest is one of an infinite class of possible nests, that is to say constructions from which any of (say) five eggs has a low probability of falling in the next month. If there is a hurricane they probably will fall, but hurricanes are rare. And so on. The goal is never exactly reached in the mathematical sense. No bird's nest is 100 per cent safe. It is often safe enough for practical purposes. So in fact for the author's equations of the form $\frac{\partial y}{\partial x_r} = 0$ we should substitute something like :

$$\text{Prob.}\left\{\frac{\partial y}{\partial x_r} < \epsilon\right\} > 1 - \eta,$$

where ϵ and η are small numbers. But this is a statistical approach. When it is adopted we can then employ the methods of cybernetics, which the author ignores, to specify how efficiently the goal is reached.

This criticism would appear carping but for the very superior attitude which the author adopts to his predecessors, whose writings he characterises as ' vague and obscure,' while promising to describe the connotation of the word life ' in exact scientific terms, i.e. in terms of clearly defined spatio-temporal relationships between quantitative variables.'

On the basis of his definition of directive correlation the author embarks on a very wide range of topics, including the nature of biological progress, not to mention religion and aesthetics. The author identifies biological progress with an increase in directive correlation ; but as he gives no quantitative method of assessing this in different cases his method is not applicable to the question whether, say, a bat is more advanced than a rabbit. I think a quantitative assessment may ultimately be possible in terms of the Shannon-Wiener theory of communication or ' messages.' The amount of information which an animal can receive in a given time is the binary logarithm of the number of classes of stimuli each of which will result in a distinct kind of behaviour in the immediate or remote future. Since nervous systems are built of components with all or none activity this number appears to be finite, and it is in fact, I believe, possible to set an upper bound to it. On these lines animals can be roughly graded.

A further criticism is as follows. In the case of truly purposive action, for example catching a train, I have a mental picture or a symbolic representation, perhaps in imagined but unspoken words, of a future event which I judge probable, namely the departure of the train at 10.40 p.m. Either there is an unbridgeable gulf between such action and a suitable coordinated reflex, in which case we had better abandon any naturalistic account of evolution, or the one shades into the other. If in a higher animal the response to a subliminal stimulus includes a subliminal cerebral activity which, if more intense, would lead to action, but whose psychical correlate is wishing rather than willing, we can see how such a gradation is possible. The author does not seriously discuss the question of such a transition, and therefore seems to the reviewer unduly to emphasise the gap between purposive and quasi-purposive activity.

To sum up, the author has put an important question. He believes that he has answered it. I venture to doubt this. On the other hand his answer is sufficiently definite to admit of constructive criticism. I have tried to give such criticism from the point of view of a statistician. Doubtless other, and more fundamental, criticisms are possible and desirable. But when all criticisms have been made, I think some positive and original contribution to the problem will remain. And this is fairly high praise for a book on philosophical biology. But I hope that if the author continues his studies, he will realise on the one hand that, since no responses are 100 per cent efficient, he cannot avoid the notion of probability, and on the other, that in modern mathematics inequalities often afford a deeper insight than equalities.

J. B. S. Haldane

REVIEWS

Problèmes de Philosophie des Sciences (*Premier Symposium—Bruxelles*, 1947), *Archives de l'Institut International des Sciences Théoretiques, Série A, Bulletin de l'Académie Internationale de Philosophie des Sciences* (*Actualités Scientifiques et Industrielles*) Hermann et C^ie^, Paris, 1948.

I *Les Methodes de la Connaissance, Actualité* 1059
  S. I. Dockx, 'De la Nature et du But de l'Institut International des Sciences Théoriques';
  H. J. Pos, 'Le Symbolisme de la Connaissance et l'Idee de l'Unité du Savoir';
  E. W. Beth, 'La Cosmologie, dite naturelle, et les Sciences mathématiques de la Nature';
  H. D. Dubarle, 'La Synthese inductive.'
II *Les Sciences et le Réel, Actualité* 1061
  J. Daujat, 'La Representation intelligible et le Réel';
  F. Gonseth, 'Les Conceptions Mathématiques et le Réel';
  Mme P. Destouches-Fevrier, 'Le Réel et la Théorie physique.'
III *Théories nouvelles de Relativité, Actualité* 1065
  E. A. Milne, 'Kinematic Relativity';
  H. Dingle, 'The Function of Time Measurement in modern Physics';
IV *Problèmes de Connaissance en Physiques moderne, Actualité* 1066
  J. L. Destouches, 'Determinisme et indeterminisme en physique moderne';
  L. de Broglie, 'L'individualité dans le monde physique';
  F. Fiala, 'Structure formelle et signification extérieure de la notion de symetrie.'
V *Problèmes de Bio-Philosophie, Actualité* 1067
  D. H. Salman, 'La Nature du Vivant';
  L. Cuénot, 'La Finalité en Biologie.'
VI *L'Unite formelle et physiologique du Vivant, Actualité* 1068
  A. Dalcq, 'La pensée moderne devant le problème de la forme';
  R. Collin, 'Les regulations hormonales et le physiologie du vivant.'
VII *Vers une Synthese moderne du Savoir, Actualité* 1109
  S. I. Dockx, 'Vers une Synthese moderne du Savoir.'
  (Including also Supplement containing 30 pp. of book reviews.)

These seven volumes, purchasable separately from Hermann et C^ie^, form the Proceedings of a Symposium on Problems of the Philosophy of Science, held in 1947 in Brussels by the International Institute of Theoretical Sciences. Each contains the papers read at one section of the Symposium, together with a report of the discussions. The series can be strongly recommended as an expert and easily-read survey of many of the problems in the philosophy of science which are of greatest interest today, though the individual volumes vary considerably in value. Some of the discussions are particularly useful and include remarks by personalities such as Weyl, Brouwer, De Donder, Le Maitre, etc., in addition to the main speakers.

I Dockx, the Director of the Institute, opens the conference with a statement of the aim of the Institute as the furthering of a new synthesis of knowledge through co-operative discussions. Pos emphasises the tension between speculative thought and rational analysis, and suggests that formal mathematical systems have to be supported by primary intuitions. Bcch discusses the relationships of different views of nature and concludes that every physical theory expresses the intuitive vision of nature of its creator. Dubarle, in an interesting paper, suggests that the main problem in the philosophy of physics lies not in its deductive aspect but in its origin and growth. He finds the crucial factor in the process of ' inductive synthesis ' by which more general principles are reached.

II This volume is of less general interest than some of the others, perhaps because, as the Chairman remarked, the problem posed may not be ripe. Daujat suggests that the intelligible is itself part of ' the real.' Gonseth attempts a reconciliation between empiricism and idealism. Paulette Destouches-Fevrier develops a systematic view of the real in quantum physics, commenting on Heisenberg's recent attempt to renounce the use of Hamiltonians. The discussions leave the reader with a vivid sense of the elusiveness of ' the real.'

III This interesting volume contains 20 pages by Milne on his Kinematic Relativity and the same by Dingle on time measurement and on his ' thermal relativity theory,' followed by a discussion with contributions by Milne, Weyl, Le Maitre, etc. Weyl and Dingle repeat a well-known criticism of Milne's theory : that it is incomplete, and cannot lead either to the ordinary time scale of physics or to the Lorentz transformations without using in addition a material rod, clock, or the equivalent. Milne's adoption of the aim of a comprehensive deductive theory appears to have blinded him, as a similar aim did Eddington, to the extent to which his theory depended on empirical observations. For he says, ' Kinematical relativity . . . aims at investigating the large-scale problems of the universe by studying the universe as it is, instead of building upwards from small-scale phenomena.' Yet his theory is based on an interpretation of a wave-length shift. and every cosmological and astrophysical theory must necessarily be built on the small-scale properties of matter and radiation which are involved in all cosmological observations. Dingle's tentative discussion of the measurement of time and of motion raises important issues.

There is no doubt of the value of these speculative investigations, but the more ambitious quasi-deductive theories of this century which rest on a rather narrow empirical basis (such as Milne's Kinematic Theory, Eddington's Fundamental Theory, and possibly even Einstein's General Theory) may eventually be modified or replaced by a comprehensive theory of the relations of the atomic, gravitational, and cosmological constants.

IV This is also valuable for those interested in the philosophy of physics. Destouches develops an orthodox interpretation of quantum indeterminacy, and goes on to define and discuss ' theories objectivistes ' and ' subjectivistes,' concluding that any theory which synthesises two contradictory theories must be more ' subjectiviste ' than they are. The future of physical theory must therefore lie in increasingly ' subjectiviste ' direction. This leads Weyl to express the timely warning that quantum theory is not in an entirely satisfactory state, and that caution is necessary in drawing such far-reaching conclusions. Destouches reaffirms the essential doctrine that correct prediction is not the sole aim of physical theory, which must also seek simplicity, i.e. aim at covering the greatest number of consequences with the minimum of hypotheses. De Broglie speculates on individuality, interactions, potential energy, and system properties in a suggestive but tentative paper. Fiala discusses various meanings of symmetry in mathematics and physics, referring particularly to Albert Lautmann's ideas (*Symmétrie et Dissymmétrie, Actualité* 1012, 1946).

V Salman considers the levels of organisation in living systems and proposes a definition of the ' living ' as a mechanism possessing a structure which is stationary, metabolic and perpetual. But this neglects the developmental aspect, which may be theoretically more important than the stationary aspect. The discussion contains important remarks by Dalcq, Cuénot (†), Prigogine (defining a *stationary state* as one in which the intensity variables are approximately constant through time), and by Weyl (on entropy).

Cuénot presents evidence for the recognition, in organisms and in their evolution, of a process of self-arrangement or organisation of a finalistic character, in addition to the established laws of physics, physical chemistry, and of chance. In discussion Salman suggests that mechanical chance in relation to particular processes may be combined with a finality and regularity of structure in the system as a whole. Two mistaken criticisms are made of Cuénot's ideas ; first, that mechanistic and finalistic expressions are mathematically equivalent, since differential equations can be put in variational form (Beth). But variational principles are not finalistic in Cuénot's sense. Secondly, that Cuénot's ideas are merely philosophical and not scientific (Wolvekamp). This view is premature and probably wrong, for the principle of decreasing asymmetry (Curie, Sellerio, Renaud, Whyte) defines a one-way process of a ' finalistic ' character capable of mathematical and empirical study.

VI Dalcq's paper is a brilliant sketch of the present knowledge of animal morphogenesis, based partly on his own researches and views, and leading to many fertile suggestions. The reviewer regards this as one of the most advanced essays in biological philosophy. Important suggestions are made concerning the linking of morphogenesis and subsequent function, and the parallel between the formative tendencies in evolution, in ontogenesis, and in human thought. The discussion fails to rise to the high level of the paper. Collin's study of hormone regulation refers to a recent development in medicine of considerable philosophical interest : Selye's ' general adaptation syndrome ' (see *British Medical Journal*, 1949, **2**, pp. 1129, 1159).

VII This consists of a general survey of the Conference and of the problem of the synthesis of modern knowledge by Dockx, and a useful 30 page supplement of short reviews of books on philosophy and science published 1947-9.

The pursuit of the philosophy of science can take two main forms : the historical and logical analysis of established scientific methods, these being regarded as definite and successful, and the constructive study of scientific methods in order that they may be improved by deeper understanding. The first is the objective study of a clearly defined subject matter, but suffers the risk that it may take too seriously special features of recently established methods (e.g. of General Relativity, or of Quantum Mechanics) which are still imperfect or require unification in a wider theory. The second is an expression of the scientific aim of reaching simpler and more comprehensive theories, and is a subtler and perhaps more difficult task, but is truly philosophical, since it recognises that science is incomplete and is continually advancing towards more powerful methods. The scientific philosopher and the philosophical scientist are both necessary, and their activities overlap and reinforce each other, but the latter can achieve the more definite advances in understanding. Judged by this criterion the papers by Dalcq and Cuénot stand out because they contain germs of ideas which may prove of constructive value to science. But the fertile interaction of the two points of view can be followed in many, if not all, of the discussions. If some of the philosophical judgments in these volumes seem doubtful, and in the coming years actually prove to be incorrect, it will be because the latest physical theory has been mistaken for the last and philosophical truth has been sought in a still developing science.

L. L. Whyte

In Defence of Philosophy, Maurice Cornforth, Lawrence & Wishart, Ltd., London, 1950. Pp. xv + 260. 12s. 6d.

The title is misleading. But the subtitle ' Against Positivism and Pragmatism ' is more informative ; for the book in fact consists of polemics against contemporary philosophic tendencies in Britain and the U.S.A., and a manifesto of the standpoint of dialectical materialism.

The critical sections of the book are all forced into the matrix of a sort of philosophical Zoroastrianism. Lenin (quoted p. 240) wrote that ' The contending parties essentially, though concealed by a pseudo-erudite quackery of new terms or by a feeble minded non-partisanship, are materialism and idealism.' Mr Cornforth therefore presents all philosophy as part of a vast cosmic conflict between, on the one side, the materialist Ormuzd and the hosts of light and progress, and on the other side, the idealist Ahriman and the polymorphous hordes of darkness and reaction. The materialist asserts and the idealist denies the existence of a (knowable) ' objective reality . . . outside our minds ' (quoted p. 144). Many positivists and pragmatists have, indeed, in incautious and aberrant moments, been misguided enough to invite the mocking onslaughts of Professor G. E. Moore by saying or implying—to adapt the words of the familiar Limerick—that the tree does not continue to be when there is no one about in the Quad (or at least that they do not *really know* that it does). These deplorable lapses into absurdity have unfortunately been sufficiently common to provide Mr Cornforth with a lot of material for his case that contemporary positivism and pragmatism are idealist, in the sense explained. Unfortunately this clash between the obvious and the preposterous is taken to be central and fundamental ; and all other aspects of contemporary philosophy are peremptorily dismissed as ' *verbal trickery*, which evades the issues of the real relations of things by scholastic quibbling about words ' (p. 25). But in spite of his self-imposed limitations the author does manage to drive some critical shafts home : for instance he deals faithfully with Dewey, a thinker much respected on the other side of the Atlantic, who ' *persistently confuses the way in which ideas are tested with what it is that is thereby tested* ' (p. 185) and whose ' elementary confusion is only possible because of the extraordinary verbosity with which he surrounds the whole discussion ' (p. 184).

The constructive part of the book expounds dialectical materialism. It is remarkable only for the appearance of the names of Zhdanov and Mao Tse-Tung among the authorities, and for the omission of all reference to *theses*, *antitheses*, and *syntheses*. Is this particular piece of Hegelian ideological lumber being quietly and belatedly junked ?

In short this is a new and revised version of Lenin's *Materialism and Empirio-Criticism*, in which the names of Dewey and Carnap replace those of Mach and Avenarius. Mr Cornforth seems to have fulfilled his task well—his reading for instance, though still not including the vital periodical literature, has been much wider than that of a well-known previous pamphleteer against positivism. If this task was to be done it was doubtless better that it should be done thus well. But how much better still had it never been done at all.

ANTONY FLEW

Beyond Realism and Idealism, W. M. Urban, George Allen & Unwin Ltd., London, 1949. Pp. 266. 18s.

THIS book is the work of one who has spent long years not only in the study of philosophy as such, but has also had occasion to reflect deeply upon the general question of values. In addition, Professor Urban has not omitted to examine the implications of physical science as they unfold in our own particular intellectual climate. The result is an essay of distinction—firstly in laying bare the accepted tenets of realism and idealism, and secondly in attempting a transcendence in which a kind of partnership may replace strife.

Rather more than half the volume is dialectical in character ; the rest concerns the actual presuppositions of cognition in the ' exact ' and ' inexact ' sciences. The conclusion is reached that science (widely interpreted) must contain both realism and idealism. Finally, a *philosophia perennis* is discussed which has much in common with that traditional discipline, though this sharing is more in the spirit than in the letter.

So much for the general trend. The author's detailed analysis of the essential nature of modern physics brings out its idealistic tendency : by contrast, recent progress in the social sciences inclines towards naturalism and realism. Thus there are two poles around which idealism and realism revolve, namely mind and matter. As Professor Urban remarks, the elusiveness of the former is underlined by the co-existence of such theories as those of *Gestalt* and Behaviourism. It is hard to credit that they both consider themselves to be ' forms ' of psychology. Compared with this, even twentieth-century ' matter ' is a little more tangible.

Fundamentally, the upshot seems to be that in effect a transcendental ' motif ' is to be enthroned, sitting majestically above the *Sturm und Drang* of partisan views. The urge that ' we must know ' (as Lossky put it) accounts for this desire. But a condition of stability for a constitutional monarchy is widely agreed to be that ' the Crown ' must not be coerced into making decisions. Furthermore, those who might be tempted to force such an issue should be reminded that their tactics may well endanger the whole system, and expose it to dire peril of collapse. This is, by analogy, both the strength and the weakness of Professor Urban's position. So long as his " Beyond Realism and Idealism " presides in kingly state as the friend and confidant of both schools, and keeps scrupulously aloof from taking sides, all will be well. The risk is that one day, one or the other will say ' Come over and help us,' and then the implicit assumption that the Crown can do no wrong will vanish into thin air, and with it much, if not all, of the sweet reasonableness of the *modus vivendi*.

Nevertheless, in spite of being somewhat repetitious, those pages provide much stimulating material for thought. Here and there are some noble passages of prose, none perhaps more charged with feeling than those grieving over the decline of metaphysics in recent years. Yet signs are not altogether wanting that, here as elsewhere, the sky is darkest before the dawn.

IAN RAWLINS

The Nature of Physical Reality, H. Margenau, McGraw-Hill Book Company, New York and London, 1950. Pp. xiii + 479. 55s. 6d.

THIS book is intended to give a philosophical interpretation of modern physical theory. The first six chapters provide the background of discussion and are concerned with the problem of perception, the character of concepts, and empirical confirmation. A chapter on Space and Time which follows is written from the viewpoint of relativity. Then the concepts of system, state, and observable are introduced. The ideas of classical physics are represented by a discussion of discrete systems, of continua, and of thermodynamics. The role of definitions in science and of probability is then examined. The discussion of statistical and of quantum mechanics follows, with separate chapters on Uncertainty and on the Pauli Principle. The ideas of modern physics (in the narrower sense) are dealt with more generally in chapters on Causality and on Models in physical theory. The book ends with an argument about what the author calls ' other kinds of reality.'

As already shown by the title of the book the exposition is kept within the frame of a somewhat antiquated and obscure theory of knowledge. Certain ideas of empiricism (i.e. phenomenalism) are blended with Kantian metaphysics. This puts severe limitations to the analysis offered. The pseudo-problem of Reality is a constant preoccupation of the author, and this could have been easily avoided if, instead of epistemology, the concepts of modern logic and semantics had been used. The very discursive style also makes for difficult reading. However, the larger part of the book is concerned with physics, and there the discussion is often interesting and informative. Although points of controversy may be raised (particularly in quantum mechanics), the book provides a competent and comprehensive survey of the main ideas of physics.

E. H. HUTTEN

BOOKS RECEIVED FOR REVIEW

Inclusion of books in this list does not preclude their being reviewed in later issues

G. Bachelard, *Etudes de Philosophie des Sciences*, Editions du Griffon, Neuchatel, 1950, pp. 174.

Collected by R. B. Bean, Edited by W. B. Bean, *Sir William Osler: Aphorisms from his Bedside Teachings and Writings*, Henry Schuman, Inc., New York, 1950, pp. 159, $2.50.

Edited by Sir Francis Darwin, *Charles Darwin's Autobiography*, Henry Schuman, Inc., New York, 1950, pp. 266, $3.50.

H. Dingle, *A Century of Science 1851–1951*, Hutchinson & Co. (Publishers) Ltd., London, 1951, pp. ix + 338, 15s.

René Dugas, *Histoire de la Mécanique*, Editions du Griffon, Neuchatel, 1950, pp. 650 + 116 diagrams, £5 15s. 6d.

A. G. N. Flew, *Logic and Language*, Basil Blackwell, Oxford, 1951, pp. viii + 206, 16s.

Ivor B. Hart, *James Watt and the History of Steam Power*, Henry Schuman, Inc., New York, 1949, pp. viii + 250, $4.00.

Otto Janssen, *Seinsordnung und Gehalt der Idealitäten*, Westkulturverlag Anton Hain, Meisenheim/Glan, 1950, pp. 284, DM. 15.50.

The late Sir James Jeans, *The Growth of Physical Science*, Cambridge University Press, 1951 (second edition), pp. x + 364, 15s.

Fr. Vinding Kruse, *The Community of the Future*, Oxford University Press (Einer Munksgaard, Copenhagen), 1950, pp. vii + 828, 42s.

Corliss Lamont, *Humanism as a Philosophy*, Philosophical Library, Inc., New York, 1950 (third edition), pp. 368, $3.75.

Rudolf Magnus, *Goethe as a Scientist*, Henry Schuman, Inc., New York, 1949, pp. xix + 259, $3.50.

Sir C. V. Raman, *The New Physics*, Philosophical Library, Inc., New York, 1951, pp. x + 144, $3.75.

Eberhard Rogge, *Axiomatik Alles Möglichen Philosophierens*, Westkulturverlag Anton Hain, Meisenheim/Glan, 1950, pp. xv + 240, DM. 13.50.

Viscount Samuel, *Essay in Physics*, Basil Blackwell, Oxford, 1951, pp. 154, 7s. 6d.

G. Sarton, *The Life of Science*, Henry Schuman, Inc., New York, 1948, pp. vii + 197, $3.50.

Dorothea Waley Singer, *Giordano Bruno*, Henry Schuman, Inc., New York (London : Constable & Co. Ltd.), 1950, pp. xi + 389, 40s.

W. M. Smart, *The Origin of the Earth*, Cambridge University Press, 1951, pp. vi + 239, 12s. 6d.

Alfred Still, *Borderlands of Science*, Philosophical Library, Inc., New York, 1950, pp. 424, $3.75.

J. L. Synge, *Science : Sense and Nonsense*, Jonathan Cape, Ltd., London, 1951, pp. 156, 8s. 6d.

F. Sherwood Taylor, *The Alchemists : Founders of Modern Chemistry*, Henry Schuman, Inc., New York, 1949, pp. x + 246, $4.00.

Rudolf Wittkower, *Architectural Principles in the Age of Humanism*, The Warburg Institute, London, 1949, pp. xi + 144, 35s.

NEWS AND NOTES

Royal Institute of Philosophy: The attention of readers is drawn to the following letter which appeared in *The Times* on 23rd January, 1951.

SIR,

The turn of the half-century has led many people to consider afresh the present state and current trends of our modern civilization. It is generally agreed that its strength lies in the striking development of science and in the advance in technology, its weakness in the lack of effective moral and spiritual standards. The intellectual world is a tangled jumble of religious, scientific and philosophic ideas, unco-ordinated and often incompatible with one another. Until this confusion is disentangled, until some kind of pattern is set on which our civilization may order itself, it cannot know what ends it should seek and how it should try to reach them.

It would seem that the duty of attacking this confusion, of trying to offer a pattern, must rest with Philosophy—but a Philosophy which shall include the valid ideas of Religion and Science. To enable philosophers to approach this task they must be acquainted with each other's work, and must pass over the boundaries between specialized departments and between national frontiers. They must have a common platform and a common audience.

To contribute to this purpose was the main object for which the British Institute of Philosophy was founded twenty-five years ago; this letter is written in order to draw the attention of a wider public to its work. That the institute had reached maturity and was representative of British philosophy in general was recognized three years ago by the conferment of the title "Royal". Its membership has now risen, from 1000 at the end of the war, to 1600. Its quarterly journal *Philosophy* enjoys a wide circulation here and abroad. But the membership is still far too small. If it were doubled the Journal could be enlarged and improved and become more influential; the lecture courses would serve a wider circle; the number of local branches could be increased. The subscription is 30s. annually, or one guinea for those who enter into a seven-year covenant; it includes receipt of the Journal; there is no entrance fee.

We invite those men and women who realize that in these times it is essential above all things to build up a body of sound constructive thought in fundamentals, as a basis for right action in morals, in politics and in economics, and who are willing to co-operate, to join the Royal Institute of Philosophy, to read its Journal and participate in its work. Application for further information and forms of membership should be made to the Honorary Secretary of the Institute, University Hall, 14 Gordon Square, London, W.C. 1.

Yours faithfully,

(Signed) Samuel (*President*)
W. D. Ross (*Chairman*)
Lindsay of Birker (*Deputy Chairman*)

Exhibition on 'Growth and Form': The Institute of Contemporary Arts is arranging to hold an exhibition on 'Growth and Form, the Development of Natural Shapes and Structures' in their gallery at 17-18 Dover Street, London, W. 1, from 3rd July to 1st September. The exhibition will include models, photographs, and films designed to illustrate the range and beauty of the visual forms of contemporary science, and is partly intended as a tribute to the late D'Arcy Wentworth Thompson and his work *Growth and Form.*

The I.C.A. also intend at the same time to publish a symposium book on 'Aspects of Growth and Form' written mainly by scientists.

The Philosophical Society of England: On Wednesday, 11th July, 1951, at 7.15 p.m. in Room 15 at the Caxton Hall (adjoining St. James's Park Station), London, Dr. A. C. Crombie will read a paper on 'Some Historical Relations between Science and Metaphysics'.

The Review of Metaphysics: The generosity of a number of American scholars has made it possible for the *Review* to continue and to expand its offer to make free subscriptions available to libraries not now in a position to pay for it, and in addition to accept a limited number of individual subscriptions at a cost of 14s., or one-third the usual rate. Complete sets of back issues are in short supply and are available only to libraries at the usual rate of £3 11s. 6d. ($10.00) per volume.

Those interested in the regular receipt of the *Review* should communicate with the Managing Editor at 201 Linsly Hall, Yale University, New Haven, Connecticut.